A.A. Samarski, E.S. Nikolaev

MÉTODOS DE SOLUCÍON DE LAS ECUACIONES RETICULARES

Tomo II

Editorial Mir Moscú

ИЗДАТЕЛЬСТВО
МИР

А. А. Самарский, Е. С. Николаев
Методы решения сеточных уравнений
Том II

Издательство «Наука»

A.A. Samarski
E.S. Nikolaev

Métodos de solución de las ecuaciones reticulares

Tomo
II

Editorial Mir
Moscú

Traducido del ruso por
Andrés Fraguela Collar
Candidato a Doctor en Ciencias
Fisicomatemáticas

Impreso en la URSS

A NUESTROS LECTORES:

Mir edita libros soviéticos traducidos al español, inglés, francés, árabe y otros idiomas extranjeros. Entre ellos figuran las mejores obras de las distintas ramas de la ciencia y la técnica: manuales para los centros de enseñanza superior y escuelas tecnológicas; literatura sobre ciencias naturales y médicas. También se incluyen monografías, libros de divulgación científica y ciencia-ficción. Dirijan sus opiniones a la Editorial Mir, 1 Rizhski per., 2, 129820, Moscú, I-110, GSP, URSS.

© Издательство «Наука
© Traducido al español. Editorial Mir 1983

Contenido

Capítulo VI. Métodos iterativos de dos capas 1

§ 1. Planteamiento del problema sobre la elección de los parámetros iterativos 11

1. Familia inicial de esquemas iterativos (11). 2. Problema para el error (12). 3. Caso autoconjugado (14)

§ 2. Método de Chebishev de dos capas 15

1. Construcción del conjunto de los parámetros iterativos (15). 2. Sobre la inmejorabilidad de la estimación apriorística (18). 3. Ejemplos de elección del operador D (19). 4. Sobre la estabilidad computacional del método (22). 5. Construcción de una sucesión óptima de los parámetros iterativos (28)

§ 3. Método de iteración simple 34

1. Elección del parámetro iterativo (33). 2. Estimación de la norma del operador de transición (35)

4. Caso no autoconjugado. Método de iteración simple . . . 37

1. Planteamiento del problema (37). 2. Minimización de la norma del operador de transición (38). 3. Minimización de la norma del operador de resolución (45). 4. Método de simetrización de una ecuación (50).

§ 5. Ejemplos de aplicación de los métodos iterativos 50

1. Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo (50). 2. Problema de Dirichlet de diferencias para la ecuación de Poisson en una región arbitraria (55). 3. Problema de Dirichlet de diferencias para una ecuación elíptica con coeficientes variables (62). 4. Problema de Dirichlet de diferencias para una ecuación elíptica con derivada mixta (69)

Capítulo VII. Métodos iterativos de tres capas 72

§ 1. Estimación de la velocidad de convergencia 72

1. Familia inicial de esquemas iterativos (72). 2. Estimación de la norma del error (73).

§ 2. Método semi-iterativo de Chebishev 76

1. Fórmulas para los parámetros iterativos (76). 2. Ejemplos de elección del operador D (79). 3. Algoritmo del método (80).

§ 3. Método de tres capas estacionario 80

1. Elección de los parámetros iterativos (80). 2. Estimación de la velocidad de convergencia (81).

§ 4. Estabilidad de los métodos de dos y tres capas respecto a los datos a priori 84

1. Planteamiento del problema (84). 2. Estimaciones de la velocidad de convergencia de los métodos (86).

Capítulo VIII. Métodos iterativos de tipo variacional 92

§ 1. Métodos gradientales de dos capas 92

1. Planteamiento del problema sobre la elección de los parámetros iterativos (92). 2. Fórmula para los parámetros iterativos (95). 3. Estimación de la velocidad de convergencia (96). 4. Inmejorabilidad de la estimación en el caso autoconjugado (98). 5. Propiedad asintótica de los métodos gradientales en el caso autoconjugado (101).

§ 2. Ejemplos de métodos gradientales de dos capas 104

1. Método del descenso más rápido (104). 2. Método de los defectos mínimos (106). 3. Método de las correcciones mínimas (107). 4. Método de los errores mínimos (108). 5. Ejemplo de aplicación de los métodos de dos capas (109).

§ 3. Método de tres capas de las direcciones conjugadas . . . 112

1. Planteamiento del problema sobre la elección de los parámetros iterativos. Estimación de la velocidad de convergencia (112). 2. Fórmulas para los parámetros iterativos. Esquema iterativo de tres capas (115). 3. Variantes de las fórmulas de cálculo (121).

§ 4. Ejemplos de métodos de tres capas 122

1. Casos particulares de los métodos de las direcciones conjugadas. (122). 2. Métodos de tres capas localmente óptimos (124).

§ 5. Aceleración de la convergencia de los métodos de dos capas en el caso autoconjugado 129

1. Algoritmo del proceso de aceleración (129). 2. Estimación de la efectividad (131). 3. Ejemplo (133).

Capítulo IX. Métodos iterativos triangulares 136

§ 1. Método de Seidel 136

1. Esquema iterativo del método (136). 2. Ejemplos de aplicación del método (141). 3. Condiciones de convergencia suficientes (144).

§ 2. Método de relajación superior 147

1. Esquema iterativo. Condiciones de convergencia suficientes (147). 2. Planteamiento del problema sobre la elección del parámetro iterativo (149). 3. Estimación del radio espectral (152). 4. Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo (155). 5. Problema de Dirichlet de diferencias para una ecuación elíptica de coeficientes variables (160).

§ 3. Métodos triangulares 163

1. Esquema iterativo (163). 2. Estimación de la velocidad de convergencia (165). 3. Elección del parámetro iterativo (167). 4. Estimación de la velocidad de convergencia de los métodos de Seidel y de relajación (168).

Capítulo X. Método alternativo triangular 172

§ 1. Teoría general del método 172

1. Esquema iterativo (172). 2. Elección de los parámetros iterativos (174). 3. Método para encontrar las magnitudes iniciales δ y Δ (178). 4. Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo (181).

§ 2. Problema de contorno de diferencias para las ecuaciones elípticas en un rectángulo 189

1. Problema de Dirichlet para una ecuación de coeficientet variables (189). 2. Método alternativo triangular modificado. (192). 3. Comparación de las variantes del método (200). 4. Tercer problema de contorno (201). 5. Problema de Dirichlet de diferencias para una ecuación de derivadas mixtas (203).

§ 3. Método alternativo triangular para ecuaciones elípticas en una región arbitraria 207

1. Planteamiento del problema de diferencia (207). 2. Construcción del método alternativo triangular (209). 3. Problema de Dirichlet para la ecuación de Poisson en una región arbitraria (214).

Capítulo XI. Método de las direcciones variables 218

§ 1. Método de las direcciones variables en el caso conmutativo 218

1. Esquema iterativo del método (218). 2. Planteamiento del problema sobre la elección de los parámetros (220). 3. Transformación fraccional lineal (223). 4. Conjunto óptimo de parámetros (225).

§ 2. Ejemplos de aplicación del método 229

1. Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo (229). 2. Tercer problema de contorno para una ecuación elíptica con variables separables (233). 3. Problema de Dirichlet de diferencias con orden de exactitud elevado (239).

§ 3. Método de las direcciones variables en el caso general . . 244

1. Caso de operadores no conmutativos (244). 2. Problema de Dirichlet de diferencias para una ecuación elíptica de coeficientes variables (247).

Capítulo XII. Métodos de resolución de ecuaciones con operadores sin signo definido y degenerados 251

§ 1. Ecuaciones con un operador real sin signo definido . . . 251

1. Esquema iterativo. Problema de la elección de los parámetros iterativos (251). 2. Transformación del operador en el caso autoconjugado (254). 3. Método iterativo con los parámetros de Chebishev (258). 4. Métodos iterativos de tipo variacional (262). 5. Ejemplos (264).

§ 2. Ecuaciones con un operador complejo 267

1. Método de iteración simple (267). 2. Método de las direcciones variables (271).

§ 3. Métodos iterativos generales para ecuaciones con un operador degenerado 275

1. Esquemas iterativos en el caso del operador no degenerado B (275). 2. Método iterativo de los defectos mínimos (279). 3. Método con los parámetros de Chébishev (283).

§ 4. Métodos especiales 289

1. Problema de Neumann de diferencias para la ecuación de Poisson en un rectángulo (289). 2. Método directo para el problema de Neumann (293). 3. Esquemas iterativos con un operador degenerado B (296).

Capítulo XIII. Métodos iterativos de resolución de ecuaciones no lineales 302

§ 1. Métodos iterativos. Teoría general 302

1. Método de iteración simple para ecuaciones con un operador monótono (302). 2. Métodos iterativos para el caso de un operador diferencial (305). 3. Método de Newton-Kantorovich (309). 4. Métodos iterativos bietápicos (313). 5. Otros métodos iterativos (317).

§ 2. Métodos de resolución de esquemas de diferencias no lineales 320

1. Esquema de diferencias para una ecuación cuasilineal elíptica unidimensional (320). 2. Método de iteración simple (329). 3. Métodos iterativos para ecuaciones elípticas cuasilineales en diferencias en un rectángulo (331). 4. Métodos iterativos para ecuaciones débilmente no lineales (338).

Capítulo XIV. Ejemplos de resolución de ecuaciones elípticas reticulares 341

§ 1. Procedimientos de construcción de esquemas iterativos implícitos 341

1. Principio de regularización en la teoría general de los métodos iterativos (341). 2. Esquemas iterativos del operador factorizado (346). 3. Procedimiento de inversión implícita del operador B (método bietápico) (351).

§ 2. Ejemplos de resolución de los problemas de contorno elípticos 354

1. Métodos directos e iterativos (354). 2. Problema de Dirichlet de diferencias con elevado orden de exactitud en el caso multidimensional (361). 3. Tercer problema de contorno para la ecuación con derivadas mixtas en el rectángulo (363). 4. Métodos iterativos para resolver un problema de diferencias (368).

§ 3. Sistemas de ecuaciones elípticas 377

1. Problema de Dirichlet para un sistema de ecuaciones elípticas en un paralelepípedo p-dimensional (377). 2. Sistema de ecuaciones de la teoría de elasticidad (382).

§ 4. Métodos de resolución de las ecuaciones elípticas en las regiones no regulares . . . . . .

1. Problemas de diferencias en las regiones de forma compuesta y métodos para resolverlos (386). 2. Método de separación de regiones (387). 3. Algoritmo del método de separación de regiones (392). 4. Método de completar la región hasta del rectángulo (396).

Capítulo XV. Métodos de resolución de ecuaciones elípticas en coordenadas curvilíneas ortogonales 400

§ 1. Planteamiento de los problemas de contorno para ecuaciones diferenciales 400

1. Ecuaciones elípticas en un sistema cilíndrico de coordenadas (400). 2. Problemas de contorno para ecuaciones en el sistema cilíndrico de coordenadas (403).

§ 2. Resolución de problemas de diferencias en un sistema cilíndrico de coordenadas 408

1. Esquemas de diferencias sin derivadas mixtas en el caso de simetría axial (408). 2. Métodos directos (412). 3. Método de las direcciones variables (414). 4. Resolución de ecuaciones definidas sobre la superficie de un cilindro (419).

§ 3. Resolución de problemas de diferencias en un sistema polar de coordenadas 423

1. Esquemas de diferencias para ecuaciones en un círculo y en un anillo (423). 2. Solubilidad de los problemas de contorno de diferencias (426). 3. Principio de superposición para un problema en un círculo (429). 4. Métodos directos de resolución de ecuaciones en un círculo y en un anillo (432). 5. Método de las direcciones variables (433). 6. Resolución de problemas de diferencias en un sector angular (436). 7. Caso general de coeficientes variables (439).

Suplemento. Construcción del polinomio de menor desviación del cero 442

Bibliografía 447

Indice alfabético 448

Capítulo VI

Métodos iterativos de dos capas

En este capítulo se examinan los métodos iterativos de dos capas para resolver la ecuación operacional $Au = f$. Los parámetros iterativos se eligen utilizando la información a priori sobre los operadores del esquema iterativo. En el § 1 se plantea el problema de elegir los parámetros para un esquema de dos capas. En los §§ 2 y 3 este problema se resuelve para el caso autoconjugado. Aquí están construidos el método de Chebishev y el método de iteración simple. En el § 4 se estudian varios métodos de elegir el parámetro iterativo en el caso no autoconjugado en función del volumen de información a priori. En el § 5 se examinan algunos ejemplos de aplicación de los métodos construidos para la resolución de las ecuaciones reticulares.

§ 1. Planteamiento del problema sobre la elección de los parámetros iterativos

1. Familia inicial de esquemas iterativos. En el capítulo V fue mostrado que los problemas de contorno de diferencias para ecuaciones elípticas representan sistemas especiales de ecuaciones algebraicas, las cuales pueden interpretarse como ecuaciones operacionales de primer género

$$Au = f \tag{1}$$

en un espacio de Hilbert real H. En algunos casos particulares estos esquemas pueden ser resueltos efectivamente por medio de los métodos directos, estudiados en los capítulos I—IV. En el caso general uno de los métodos aproximados de resolución de ecuaciones elípticas reticulares es el método de las iteraciones. Comenzaremos el estudio de los métodos iterativos por los métodos de dos capas más simples: *el método de Chebishev y el método de iteración simple.*

Para la resolución aproximada de la ecuación (1) de un operador lineal A no degenerado, prefijado en H, examinaremos *el esquema iterativo de dos capas implícito*

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f,\ k=0,\ 1,\ \ldots \qquad (2)$$

con una aproximación inicial arbitraria $y_0 \in H$. Aquí $\{\tau_k\}$ es una sucesión de parámetros iterativos, y B es un operador lineal no degenerado arbitrario, que actúa en H. Respecto a la mejor elección del operador B el problema será estudiado por separado, aquí sólo señalaremos que el operador B debe ser fácilmente inversible.

Investigaremos la convergencia del esquema iterativo (2) en el espacio energético H_D, generado por un operador D en H autoconjugado y definido positivo arbitrario.

Como el operador B no es fijo, entonces (2) genera una familia de esquemas iterativos a la cual llamaremos *familia inicial.*

En el capítulo V fue mostrado, que para estudiar la convergencia de un método iterativo es necesario investigar el comportamiento de la norma en H_D del error $z_k = y_k - u$ cuando $k \to \infty$, donde y_k es la aproximación iterativa, obtenida por el esquema (2) y u es la solución de la ecuación (1). El método iterativo converge en H_D, si la norma del error z_k en H_D converge a cero, cuando k tiende a infinito.

Ya que la velocidad de convergencia depende de la elección de los parámetros iterativos τ_k, entonces ellos deben ser elegidos de manera tal, que la velocidad de convergencia sea máxima.

2. **Problema para el error.** Investiguemos primeramente la convergencia de los esquemas iterativos de dos capas (2). Para esto obtendremos la ecuación a la cual satisface el error z_k.

Sustituyendo $y_k = z_k + u$ para $k = 0,\ 1,\ \ldots$ en (2) y teniendo en cuenta la ecuación (1), hallamos

$$B\frac{z_{k+1}-z_k}{\tau_{k+1}}+Az_k=0, \quad k=0,\ 1,\ \ldots,\ z_0=y_0-u,$$

es decir el error z_k satisface una ecuación homogénea. Resolviendo esta ecuación con respecto a z_{k+1}:

$$z_{k+1} = (E - \tau_{k+1}B^{-1}A)\, z_k$$

y poniendo $z_k = D^{-1/2}x_k$, pasamos a una ecuación para el error equivalente x_k, la cual contendrá un operador. La ecuación para x_k tendrá la forma

$$x_{k+1} = S_{k+1}x_k, \quad S_{k+1} = E - \tau_{k+1}C, \quad k = 0, 1, \ldots, \tag{3}$$

donde $C = D^{1/2}B^{-1}AD^{-1/2}$. En virtud del cambio hecho es válida la igualdad

$$\| x_k \| = \| D^{1/2}z_k \| = \| z_k \|_D,$$

por eso el problema de investigar la convergencia del método iterativo (2) en H_D se reduce al estudio de la sucesión numérica $\| x_k \|$, $k = 1, 2, \ldots$, donde x_k está definido en (3).

Hallemos la solución de la ecuación (3). De (3) obtenemos

$$x_k = T_{k,0}x_0, \quad T_{k,0} = \prod_{i=1}^{k} S_i = S_nS_{n-1} \ldots S_1.$$

De aquí se infiere la siguiente estimación para la norma del error z_k en H_D:

$$\| z_k \|_D = \| x_k \| \leqslant \| T_{k,0} \| \| x_0 \| = \| T_{k,0} \| \| z_0 \|_D. \tag{4}$$

El operador $T_{k,0}$ se llama *operador de resolución (solvente)* para la k-ésima iteración, y S_k *operador de transición* de la $(k-1)$-ésima iteración.

De la estimación (4) se desprende, que el método iterativo (2) converge en H_D, si la norma del operador de resolución $T_{k,0}$ tiende a cero, al tender k a infinito.

De esta forma el problema de investigar la convergencia del esquema iterativo (2) en H_D está reducido al estudio del comportamiento de la norma del operador solvente $T_{k,0}$ en el espacio H en dependencia del número k de la iteración. El operador de resolución $T_{k,0}$ se determina por el operador C y los parámetros iterativos $\tau_1, \tau_2, \ldots, \tau_k$.

Considerando el operador C fijo, nos plantearemos el problema de elegir los parámetros $\{\tau_k\}$ de manera tal, que el método iterativo converga. Entre los métodos iterativos convergentes el método *óptimo* será, evidentemente, aquel cuyos parámetros τ_k garanticen la obtención de una exactitud dada $\varepsilon > 0$ en un número mínimo de iteraciones. En virtud de la estimación (4) a esta exigencia se le puede dar la siguiente forma equivalente: para n dado construir un

conjunto de parámetros iterativos $\tau_1, \tau_2, \ldots, \tau_n$, para la cual la norma del operador. $T_{n,0}$ sea mínima.

3. **Caso autoconjugado.** Daremos ahora un planteamiento riguroso del problema sobre la mejor elección de los parámetros iterativos para el esquema de dos capas (2). Este problema tendrá solución para determinadas suposiciones respecto de los operadores A, B y D. Formulemos estas suposiciones.

1) Supondremos, que los operadores A, B y D son tales, que el operador $DB^{-1}A$ es autoconjugado en H. Si se cumple esta suposición, entonces diremos, que estamos examinando el caso autoconjugado.

2) Sean dadas las constantes γ_1 y γ_2 de equivalencia energética entre los operadores D y $DB^{-1}A$, es decir las constantes de las desigualdades

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0,$$
$$DB^{-1}A = (DB^{-1}A)^*. \tag{5}$$

La segunda suposición determina el tipo de información a priori sobre los operadores del esquema iterativo; esta información se utiliza durante la construcción de las fórmulas para los parámetros iterativos en el caso autoconjugado. El ejemplo más simple, para el cual se cumple la suposición sobre la autoconjugación del operador $DB^{-1}A$, es el siguiente: $A = A^*$, $D = B = E$, es decir se examina un esquema explícito en el espacio de partida H para la ecuación (1) con el operador autoconjunto A. En este caso la información a priori consiste en prefijar las cotas del operador A. Ejemplos más complejos de elección del operador D serán examinados más adelante.

De esta forma, supongamos que se cumplen las condiciones (5). De (5) se deduce, que el operador $C = D^{-1/2} \times (DB^{-1/2}A) D^{1/2}$ es autoconjugado en H, y γ_1, γ_2 son sus cotas, es decir

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \gamma_1 > 0, \quad C = C^* =$$
$$= D^{-1/2} (DB^{-1}A) D^{-1/2}. \tag{6}$$

En efecto, poniendo $x = 0^{-1/2}y$ en las desigualdades

$$\gamma_1 (Dx, x) \leqslant (DB^{-1}Ax, x) \leqslant \gamma_2 (Dx, x),$$

obtendremos las desigualdades (6). De esta manera, las suposiciones hechas más arriba respecto a los operadores A, B y D son equivalentes a las condiciones (6).

Formulemos ahora el problema sobre la elección óptima de los parámetros iterativos para el esquema (2). De la definición del operador solvente $T_{k,0}$ y de las condiciones (6) se desprende, que el operador $T_{k,0} = T_{k,0}(C)$ es autoconjugado en H y la norma del polinomio operacional $T_{n,0}(C)$ se estima de la siguiente manera:

$$\| T_{n,0} \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \left| \prod_{k=1}^{n} (1 - \tau_k t) \right|.$$

De la estimación (4) se deduce, que en el caso autoconjugado los parámetros iterativos $\tau_1, \tau_2, \ldots, \tau_n$ deben ser elegidos de manera tal, que el máximo del módulo del polinomio $P_n(t) = \prod_{k=1}^{n} (1 - \tau_k t)$, construido según estos parámetros, sea mínimo en el segmento $[\gamma_1, \gamma_2]$, es decir necesitamos hallar los parámetros de la condición

$$\min_{\{\tau_k\}} \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \left| \prod_{k=1}^{n} (1 - \tau_k t) \right| = \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |p_n(t)|.$$

Entonces para el error del método (2) será cierta la estimación $\| z_n \|_D \leqslant q_n \| z_0 \|_D$, donde

$$q_n = \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |P_n(t)|.$$

El problema formulado más arriba es un problema clásico del minimax. En el § 2 se da la resolución de este problema y se construirá el conjunto de los parámetros $\tau_1, \tau_2, \ldots, \tau_n$. El método iterativo con dicho conjunto de parámetros se llama *método de Chebishev*. En la literatura este método también lo llaman *método de Richardson*.

§ 2. Método de Chebishev de dos capas

1. Construcción del conjunto de los parámetros iterativos. En el § 1 fue mostrado que la construcción de un conjunto óptimo de parámetros iterativos $\tau_1, \tau_2, \ldots, \tau_n$ se reduce a encontrar al polinomio $P_n(t)$ del tipo $P_n(t) = \prod_{k=1}^{n} (1 - \tau_k t)$, cuyo módulo máximo sobre el segmento $[\gamma_1, \gamma_2]$ sea mínimo.

Resolvemos este problema. Ya que el tipo de polinomio se determina por la condición de normación $P_n(0) = 1$, entonces el problema indicado se formula de la siguiente forma: entre todos los polinomios de grado n, que toman el valor 1 en el punto $t = 0$, hallar el polinomio que menos se desvía del cero sobre el segmento $[\gamma_1, \gamma_2]$, el cual no contiene el punto 0.

La solución de este problema fue obtenida por el matemático ruso V. A. Markov en 1892 y está citada en el suplemento. El polinomio buscado $P_n(t)$ tiene la forma

$$P_n(t) \equiv q_n T_n\left(\frac{1-\tau_0 t}{\rho_0}\right), \quad q_n = \frac{1}{T_n\left(\frac{1}{\rho_0}\right)} \tag{1}$$

donde $T_n(x)$ es el polinomio de Chebishev de primer género y de grado n,

$$T_n(x) = \begin{cases} \cos(n \arccos x), & |x| \leqslant 1, \\ \operatorname{ch}(n \operatorname{Arch} x), & |x| \geqslant 1, \end{cases}$$

$$q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}}, \quad \tau_0 = \frac{2}{\gamma_1+\gamma_2},$$

$$\rho_0 = \frac{1-\xi}{1+\xi}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}. \tag{2}$$

Con esto $\max\limits_{\gamma_1 \leqslant t \leqslant \gamma_2} |P_n(t)| = q_n$. De aquí se deduce la estimación para la norma del error z_n en H_D:

$$\| z_n \|_D \leqslant q_n \| z_0 \|_D, \tag{3}$$

donde q_n está definido en (2).

Obtengamos las fórmulas para los parámetros iterativos. Como los polinomios que se encuentran en los miembros primero y segundo de (1), toman el mismo valor, igual a 1, para $t = 0$, entonces la identidad en (1) tendrá lugar solamente en aquel caso, cuando el conjunto de las raíces de los polinomios $P_n(t)$ y $T_n \frac{1-\tau_0 t}{\rho_0}$ coincidan. El polinomio $P_n(t)$ tiene las raíces $1/\tau_k$, $k = 1, 2, \ldots, n$, y el polinomio $T_n(x)$ tiene las raíces iguales a $-\cos\left(\frac{2i-1}{2n}\pi\right)$, $i = 1, 2, \ldots, n$. Si designamos mediante $\mathfrak{M}_n$ el conjunto de raíces del polinomio de Chebishev $T_n(x)$:

$$\mathfrak{M}_n = \left\{-\cos\frac{2i-1}{2n}\pi, \; i = 1, 2, \ldots, n\right\}, \tag{4}$$

entonces obtendremos la siguiente fórmula para los parámetros iterativos:

$$\tau_k = \tau_0/(1 + \rho_0\mu_k), \quad \mu_k \in \mathfrak{M}_n, \quad k = 1, 2, \ldots, n. \tag{5}$$

Aquí $\mu_k \in \mathfrak{M}_n$ significa, que en calidad de μ_k deben elegirse sucesivamente todos los elementos del conjunto $\mathfrak{M}_n$.

De la fórmula obtenida para los parámetros τ_k se ve, que para el cálculo de los parámetros iterativos es necesario dar el número n de iteraciones. Estimemoslo. En calidad de condición para la terminación del proceso de iteraciones se toma la desigualdad

$$\| z_n \|_D \leqslant \varepsilon \| z_0 \|_D$$

y se llama *número de iteraciones* el menor entero n, para el cual se cumple esta desigualdad.

De (3) se deduce, que para el método examinado el número de iteraciones se encuentra de la desigualdad $q_n \leqslant \varepsilon$. Utilizando (2), resolveremos esta desigualdad. Obtendremos

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \ln\left(\frac{1}{\varepsilon} + \sqrt{\frac{1}{\varepsilon^2}1 - }\right)\Big/\ln\frac{1}{\rho_1}.$$

Frecuentemente se utiliza una fórmula más simple para $n_0(\varepsilon)$

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \ln\frac{2}{\varepsilon}\Big/\ln\frac{1}{\rho_1}. \tag{6}$$

Después que ha sido hallado el número exigido n de iteraciones, por la fórmula (5) se puede construir el conjunto de parámetros iterativos.

De esta forma, para el esquema implícito de dos capas

$$B\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, y_0 \in H, \tag{7}$$

está demostrado

TEOREMA 1. *Supongamos que se cumplen las condiciones*

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0, \quad DB^{-1}A = (DB^{-1}A)^*, \quad D = D^* > 0. \tag{8}$$

Entonces el proceso iterativo de Chebishev (7), (4), (5), (2) *converge en* H_D *y para el error* z_n *tiene lugar la estimación* (3). *Para el número de iteraciones es válida la estimación* (6).

De las estimaciones obtenidas se desprende, que en el caso autoconjugado la velocidad de convergencia del método

de Chebishev depende de la relación $\xi = \gamma_1/\gamma_2$, y al mismo tiempo la velocidad de convergencia será mayor mientres mayor sea ξ.

2. Sobre la inmejorabilidad de la estimación apriorística. Mostremos ahora, que en la clase de las aproximaciones iniciales arbitrarias y_0 la estimación, obtenida en el teorema 1 para el error del método de Chebishev, es inmejorable en el caso de un espacio H de dimensión finita. Es suficiente indicar una aproximación inicial y_0 con la cual para la norma del error equivalente x_k tenga lugar la igualdad $\| x_n \| = q_n \| x_0 \|$. Nosotros hallaremos el error inicial x_0, el cual garantiza el cumplimiento de esta igualdad, y en virtud de la relación entre los errores z_k y x_k, $z_k = D - \frac{1}{2} x_k$, la aproximación inicial y_0 se determina entonces por la fórmula $y_0 = u + D^{-\frac{1}{2}} x_0$.

Hallemos el x_0 buscado. Sea H un espacio de dimensión finita ($H = H_N$). Ya que el operador C es autoconjugado en H, entonces existe un sistema completo de funciones propias $v_1, v_2, \ldots, v_N$ del operador C. Designemos mediante λ_k el valor propio del operador C, correspondiente a la función propia v_k. Sean los valores propios ordenados $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_N$. Entonces como cotas del operador C se pueden tomar $\gamma_1 = \lambda_1$ $\gamma_2 = \lambda_N$.

Tomemos en calidad de error inicial x_0 la función propia v_1. De la ecuación para el error x_k:

$x_{k+1} = (E - \tau_{k+1} C)\, x_k, \quad k = 0, 1, \ldots, x_0 = v_1$ y

la igualdad $Cv_k = \lambda_k v_k$ obtendremos sucesivamente

$$x_1 = (E - \tau_1 C)\, x_0 = (1 - \tau_1 \gamma_1)\, v_1 = (1 - \tau_1 \gamma_1)\, x_0,$$

$$x_2 = (E - \tau_2 C)\, x_1 = (1 - \tau_1 \gamma_1)\, (E - \tau_2 C)\, x_0 =$$

$$= (1 - \tau_1 \gamma_1)\, (1 - \tau_2 \gamma_1)\, x_0,$$

. . . . . .

$$x_n = \prod_{k=1}^{n} (1 - \tau_k \gamma_1)\, x_0 = P_n (\gamma_1)\, (x_0).$$

Sustituyendo $t = \gamma_1$ en (1) y teniendo en cuenta la igualdad $1 - \tau_0 \gamma_1 = \rho_0$, calcularemos $P_n (\gamma_1) = q_n T_n (1) = q_n$ y, por consiguiente,

$$x_n = q_n x_0, \quad \| x_n \| = q_n \| x_0 \|,$$

lo que se exigía demostrar.

De esta forma, está mostrado, que la estimación apriorística obtenida en el teorema 1 no es mejorable en la clase de las aproximaciones iniciales arbitrarias.

3. Ejemplos de elección del operador D. Citemos algunos ejemplos de elección del operador D. Recordemos, que el método de Chebishev se examina suponiéndose la autoconjugación del operador $DB^{-1}A$. Más abajo serán indicadas las exigencias a los operadores A y B, con las cuales esta suposición se cumple para el operador D elegido. Para cada elección concreta del operador D serán expuestas las desigualdades que prefijan la información a priori sobre los operadores del esquema iterativo, Esta información se utiliza para construir el conjunto de parámetros iterativos en el método de Chebishev.

EXAMINEMOS EL PRIMER EJEMPLO. Sean A y B operadores autoconjugados y definidos positivos en H. Entonces en calidad del operador D se puede tomar uno de los siguientes operadores: A o B. Si, además, el operador B está acotado en H, entonces se puede tomar $D = AB^{-1}A$. Con esto la información a priori se reduce a prefijar las constantes de equivalencia energética de los operadores A y B:

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0, \quad B > 0. \tag{9}$$

En efecto, es necesario mostrar que se cumplen las siguientes condiciones: el operador D elegido es autoconjugado y definido positivo en H, el operador $DB^{-1}A$ es autoconjugado en H, y las desigualdades (8) y (9) son equivalentes.

La autoconjugación de los operadores D y $DB^{-1}A$ para todos los casos examinados se deduce de la autoconjugación de los operadores A y B. Para el caso, cuando $D = A$ o $D = B$, el hecho que D sea definido positivo se infiere de que los operadores A y B lo son también. Mostremos ahora, que el operador $D = AB^{-1}A$ es también definido positivo en H.

En efecto, supongamos que se cumplen las condiciones formuladas más arriba sobre los operadores A y B: $A = A^* \geqslant \alpha E$, $B = B^* \geqslant \beta E$, $\| Bx \| \leqslant M \| x \|$, $\alpha, \beta > 0$, $M < \infty$. De estas condiciones y de los lemas 6 y 8 del § 1 del cap. V obtendremos, que $B^{-1} \geqslant \frac{1}{M} E$ y $(Ax, Ax) \geqslant \alpha (Ax, x) \geqslant \alpha^2 (x, x)$. De aquí hallaremos para la energía del operador D la estimación por debajo o inferior $(Dx, x) = (AB^{-1}Ax, x) = (B^{-1}Ax, Ax) \geqslant \frac{1}{M} (Ax, Ax) \geqslant$

$\geqslant \frac{\alpha^2}{M}(x, x)$, es decir $D \geqslant \frac{\alpha^2}{M} E$. Por consiguiente, está demostrado que el operador $D = AB^{-1}A$ es definido positivo.

Mostremos ahora, que las desigualdades (8) y (9) son equivalentes para el ejemplo examinado. En efecto, supongamos que se cumplan las desigualdades (9):

$$\gamma_1 (Bx, x) \leqslant (Ax, x) \leqslant \gamma_2 (Bx, x), \quad \gamma_1 > 0. \qquad (10)$$

Si $D = B$, entonces $DB^{-1}A = A$ y, por consiguiente, las desigualdades (10) y (8) coinciden. Sea ahora $D = AB^{-1}A$. En este caso $DB^{-1}A = AB^{-1}AB^{-1}A$ y, poniendo $x = B^{-1}Ay$ en (10), obtendremos

$$\gamma_1 (AB^{-1}Ay, y) \leqslant (AB^{-1}Ay, B^{-1}Ay) \leqslant \gamma_2 (AB^{-1}Ay, y)$$

ó

$$\gamma_1 (Dy, y) \leqslant (DB^{-1}Ay, y) \leqslant \gamma_2 (Dy, y),$$

es decir, obtendremos las desigualdades (8). El paso inverso de (8) a (10) es evidente.

Sea $D = A$, entonces $DB^{-1}A = AB^{-1}A$. Del lema 9 del § 1 cap. V se deduce que para los operadores A y B autoconjugados y definidos positivos las desigualdades (10) y las desigualdades

$$\gamma_1 (A^{-1}x, x) \leqslant (B^{-1}x, x) \leqslant \gamma_2 (A^{-1}x, x), \quad \gamma_1 > 0$$

son equivalentes. Poniendo aquí $x = Ay$, obtenemos las desigualdades (8). El paso inverso es evidente.

Esta desigualdad permite demostrar inmediatamente que D es definido positivo:

$$(Dx, x) \geqslant \alpha\gamma_1 (x, x).$$

En efecto, $(Dx, x) = (B^{-1}Ax, Ax) \geqslant \gamma_1 (A^{-1}Ax, Ax) = \gamma_1 (Ax, x) \geqslant \gamma_1\alpha (x, x)$.

SEGUNDO EJEMPLO. Sean A y B operadores en H autoconjugados, definidos positivos y conmutativos: $A = A^* > 0$, $B = B^* > 0$, $AB = BA$. Si en calidad del operador D tomamos el operador A^2, entonces la información a priori puede ser dada en forma de las desigualdades (9).

Realmente la autoconjugación y la definición positiva del operador D se desprenden de la autoconjugación y no degeneración del operador A. Luego, $DB^{-1}A = A(AB^{-1})A$, y como los operadores A y B conmutan, entonces también conmutan los operadores A y B^{-1}. De aquí y de la autocon-

jugación de los operadores A y B se deduce la autoconjugación del operador $DB^{-1}A$.

En el caso dado las desigualdades (8) tienen la forma

$$\gamma_1 (Ax, Ax) \leqslant (AB^{-1}Ax, Ax) \leqslant \gamma_2 (Ax, Ax), \gamma_1 > 0.$$

Poniendo aquí $x = A^{-1}B^{1/2}y$ y utilizando la conmutatividad de la raíz del operador B con el operador A, hallamos

$$\gamma_1 (By, y) \leqslant (Ay, y) \leqslant \gamma_2 (By, y)$$

es decir obtendremos la desigualdad (9). El paso inverso de (9) a (8) es evidente.

EXAMINEMOS UN EJEMPLO MAS. Sean A y B operadores no degene os arbitrarios, que satisfacen la condición

$$B^*A = A^*B. \tag{11}$$

Si en calidad de D elegimos el operador A^*A, entonces la información a priori puede ser dada en forma de las desigualdades

$$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx) \leqslant \gamma_2 (Bx, Bx), \quad \gamma_1 > 0. \tag{12}$$

La autoconjugación del operador D es evidente, y la definición positiva se deduce de la no degeneración del operador A. Como el operador B es no degenerado, entonces las condiciones (11) pueden ser escritas en forma de las condiciones $AB^{-1} = (B^*)^{-1}A^*$, las cuales expresan la autoconjugación del operador AB^{-1}. De aquí obtenemos, que el operador $DB^{-1}A = A^*AB^{-1}A$ es autoconjugado en H. Después, poniendo $x = B^{-1}Ay$ en (12) obtendremos

$$\gamma_1 (Ay, Ay) \leqslant (AB^{-1}Ay, Ay) \leqslant \gamma_2 (Ay, Ay)$$

ó

$$\gamma_1 (Dy, y) \leqslant (DB^{-1}Ay, y) \leqslant \gamma_2 (Dy, y).$$

De este modo, de las desigualdades (12) se deducen las desigualdades (8). El paso inverso de (8) a (12) es evidente.

Como conclusión notemos, que para el caso de los operadores A y B autoconjugados, definidos positivos y acotados en H el método iterativo de Chebishev converge en H_D, donde $D = A$, B ó $AB^{-1}A$ (y si, además, A y B conmutan, entonces también para $D = A^2$), con una misma velocidad, determinada por la relación entre las constantes γ_1 y γ_2 de las desigualdades (9).

Especialmente señalemos los casos $D = AB^{-1}A$ y $D = A^*A$. Para tal elección del operador D la norma del

error en H_D puede ser calculada en el proceso de las iteraciones. En efecto, para $D = AB^{-1}A$ obtendremos

$$\| z_n \|_D^2 = (Dz_n, z_n) = (B^{-1}Az_n, Az_n) = (B^{-1}r_n, r_n) = \\ = (w_n, e_n),$$

y para $D = A^*A$:

$$\| z_n \|_D^2 = (Az_n, Az_n) = (r_n, r_n),$$

donde $r_n = Az_n = Ay_n - Au = Ay_n - f$ es el error de la n-ésima iteración, y $w_n = B^{-1}r_n$ es la corrección. Estas magnitudes pueden hallarse en el proceso de las iteraciones.

4. Sobre la estabilidad computacional del método. Estudiando la convergencia del método de Chebishev se supuso, que el proceso de cálculo es ideal, es decir que los cálculos se producen con un número infinito de signos. En el proceso computacional real todos los cálculos se realizan con un número finito de signos, y en cada etapa del cómputo aparecen errores de redondeo. Los errores de redondeo de los resultados de operaciones aritméticas generan el error computacional del método.

En los métodos iterativos el error computacional del método se forma de los errores, admitidos en cada iteración. Si el número de iteraciones es suficientemente grande, y el método iterativo posea la propiedad de acumular los errores de redondeo de cada paso iterativo, entonces el error computacional de tal método puede resultar tan grande que ocurrirá una pérdida completa de exactitud y la aproximación iterativa y_n se diferenciará fuertemente de la solución buscada. Por eso para los métodos iterativos es importante estudiar el mecanismo de aparición del error computacional y encontrar aquellas etapas del algoritmo, en las cuales tiene lugar un crecimiento del error computacional del método. En una serie de casos algunos cambios en el proceso del cálculo permiten disminuir esencialmente el crecimiento del error computacional y convertir dicho método en todo útil para su aplicación práctica.

Otra singularidad del proceso computacional real está relacionada con la presencia en la MCE de un «cero de máquina» y de un «infinito de máquina». Estos conceptos caracterizan el orden admisible de los números, que pueden ser representados en la MCE. Por ejemplo en la MCE BESM-6 en el régimen de exactitud simple pueden ser representados números reales, cuya magnitud absoluta pertenece al diapa-

són desde 10^{-19} hasta 10^{19}. Estas son las cotas para el «cero de máquina» y el «infinito de máquina». Si como resultado de los cálculos en la MCE aparece una magnitud que no pertenece a este intervalo, entonces los cálculos se interrumpen y tiene lugar la así llamada «parada de reparación» de fallo. (parep). Por eso la exigencia de un proceso iterativo «sin paradas de fallo» es natural.

De ese modo, los métodos iterativos deben ser «sin paradas de fallo» y estables respecto a los errores de redondeo.

En el punto 1 del § 2 fue construido el método de Chebishev de dos capas. En el teorema 1 está demostrado, que si se cumplen n iteraciones con los parámetros $\tau_k = \tau_0/(1 + \rho_0\mu_k)$, $\mu_k \in \mathfrak{M}_n$, $k = 1, 2, \ldots, n$, entonces para el error z_n será válida la estimación $\| z_n \|_D \leqslant q_n \| z_0 \|_D$. En calidad de μ_k se eligen sucesivamente todos los elementos del conjunto $\mathfrak{M}_n$ y al mismo tiempo el orden de elección es arbitrario.

Estudiemos la estabilidad computacional del método de Chebishev. Para más precisión vamos a considerar, que μ_k es el k-ésimo elemento del conjunto $\mathfrak{M}_n$. Entonces diferentes ordenamientos del conjunto $\mathfrak{M}_n$ generarán diferentes sucesiones $\{\mu_k\}$ y, por consiguiente, sucesiones diferentes de parámetros iterativos $\{\tau_k\}$.

Desde el punto de vista de un proceso computacional ideal todas las sucesiones de parámetros iterativos de Chebishev son equivalentes, es decir, cada sucesión debe garantizar la obtención de una misma aproximación y_n y, por lo tanto, de una exactitud dada después de cumplirse n iteraciones. La presencia de errores de redondeo en el proceso computacional real conduce a la no equivalencia de las sucesiones de parámetros iterativos.

Ilustremos esta afirmación en un ejemplo. Supongamos que sobre la red $\bar{\omega} = \{x_i = ih,\ 0 \leqslant i \leqslant N,\ h = 1/N\}$ se exige hallar la solución del siguiente problema de diferencias:

$$\Lambda y = y_{\bar{x}x} - dy = -\varphi(x),\ x \in \omega,$$
$$y(0) = 0, \quad y(1) = 1, \quad d = \text{const} > 0.$$

En el § 2 del cap. V fue mostrado, que este problema de diferencias puede ser reducido a la ecuación operacional

$$Ay = f, \tag{13}$$

en la cual el operador A se define de la siguiente forma: $Ay = -\Lambda\mathring{y}$, donde $y \in H$, $\mathring{y} \in \mathring{H}$, $\mathring{y}(x) = y(x)$ para $x \in \omega$. Aquí $\mathring{H}$ es el conjunto de funciones reticulares, definidas sobre $\overline{\omega}$ y que se anulan para $x = 0$ y $x = 1$, y H es el espacio de las funciones reticulares definidas sobre ω, con el producto escalar $(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h$. El miembro derecho f de la ecuación (13) se diferencia del miembro derecho φ del esquema de diferencias solamente en los nodos fronterizos de la red: $f(x) = \varphi(x)$, $h \leqslant x \leqslant 1 - 2h$, $f(1 - h) = \varphi(1 - h) + 1/h^2$.

Para la resolución aproximada de la ecuación (13) examinemos el método de Chebishev explícito

$$\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f,\ k = 0,\ 1,\ \ldots,\ y_0 \in H. \qquad (14)$$

Ya que los operadores A y $B = E$ son autoconjugados en H, entonces de los ejemplos examinados en el punto 3 del § 2 se deduce, que en calidad de información a priori para el método de Chebishev (14) es suficiente prefijar las cotas del operador A: $\gamma_1 E \leqslant A \leqslant \gamma_2 E$, $\gamma_1 > 0$, si como operador D tomamos el operador $B = E$. Es evidente, que γ_1 y γ_2 coinciden con los valores propios máximo y mínimo del operador de diferencias Λ, es decir

$$\gamma_1 = \frac{4}{h^2} \operatorname{sen}^2 \frac{\pi h}{2} + d,\quad \gamma_2 = \frac{4}{h^2} \cos^2 \frac{\pi h}{2} + d.$$

Los parámetros iterativos τ_k se calculan por las fórmulas

$$\tau_k = \tau_0/(1 + \rho_0 \mu_k),\quad \mu_k \in \mathfrak{M}_n,$$
$$k = 1,\ 2,\ \ldots,\ n,$$
$$\tau_0 = 2/(\gamma_1 + \gamma_2),\quad \rho_0 = (\gamma_2 - \gamma_1)/(\gamma_2 + \gamma_1). \qquad (15)$$

Se examinaron tres sucesiones de parámetros iterativos, definidos por los siguientes ordenamientos de $\mathfrak{M}_n$:

1) sucesión «directa»

$\mathfrak{M}_n = \mathfrak{M}_n^{(1)} = \{\sigma_1 \sigma_2,\ \ldots,\ \sigma_n\}$, es decir $\mu_k = \sigma_k$, $k = 1,\ 2,\ \ldots,\ n$;

2) sucesión «inversa»

$\mathfrak{M}_n = \mathfrak{M}_n^{(2)} = \{\sigma_n,\ \sigma_{n-1},\ \ldots,\ \sigma_1\}$, es decir $\mu_k = \sigma_{n-k+1}$, $k = 1,\ 2,\ \ldots,\ n$;

3) sucesión «alternada»

$$\mathfrak{M}_n = \mathfrak{M}_n^{(3)} = \{\sigma_1, \sigma_n, \sigma_2, \sigma_{n-1}, \ldots\}, \text{ es decir}$$

$$\mu_{2k-1} = \sigma_k, \quad \mu_{2k} = \sigma_{n-k+1}, \quad k = 1, 2, \ldots, n/2.$$

Aquí está designado $\sigma_k = -\cos\frac{2k-1}{2n}\pi$.

Los cálculos se realizaron de la siguiente forma: se prefijó el número n de iteraciones y por el esquema (14), (15) para cada sucesión de parámetros iterativos se realizaron n iteraciones. La exactitud real, que se alcanzó después de cumplirse n iteraciones, se determinó por la fórmula

$$\varepsilon_{\text{real}} = \frac{\|y_n - u\|}{\|y_0 - u\|}.$$

Para comparar se ha calculado el valor de q_n, donde

$$q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

el cual determina la exactitud teórica del método, cuando el número de iteraciones es igual a n. En todos los cálculos la aproximación inicial y_0 se tomó igual a cero sobre ω. La solución exacta del problema de diferencias $y(x) = x$ corresponde al segundo miembro $\varphi(x) = dx$. El coeficiente d se eligió de manera tal, que γ_1 fuera igual a 0,1:

$$\gamma_1 = 0{,}1, \quad \gamma_2 = 0{,}1 + \frac{4}{h^2}\cos\pi h, \quad \frac{1}{\xi} = \frac{40}{h^2}\cos\pi h + 1.$$

Los resultados de los cálculos para $N = 10$ se exponen en la tabla 5. En dicha tabla, además de las sucesiones indicadas de parámetros, se dan los resultados para el ordenamiento óptimo del conjunto $\mathfrak{M}_n^*$, que será descrito más abajo.

Los cálculos realizados muestran, que para el proceso computacional real las sucesiones de parámetros iterativos examinados no son realmente equivalentes. Los cálculos demostraron dos particularidades características del proceso computacional real: la posibilidad de «parep» proporcionada por el crecimiento de las soluciones iterativas intermedias y la posibilidad de determinada pérdida de exactitud en una situación sin paradas de reperación provocada por la acumulación de errores de redondeo.

El hecho que causa tal inestabilidad computacional del método para algunas sucesiones de parámetros iterativos es, que la norma del operador de transición de una iteración

Tabla 5

n	q_n	e real			
		$\mathfrak{M}_n^{(1)}$	$\mathfrak{M}_n^{(2)}$	$\mathfrak{M}_n^{(3)}$	$\mathfrak{N}_n^*$
16	$8{,}79\cdot10^{-1}$	$8{,}14\cdot10^{-1}$	$8{,}14\cdot10^{-1}$	$8{,}14\cdot10^{-1}$	$8{,}14\cdot10^{-1}$
24	$7{,}58\cdot10^{-1}$	$9{,}62\cdot10^{-1}$	$7{,}11\cdot10^{-1}$	$7{,}11\cdot10^{-1}$	$7{,}11\cdot10^{-1}$
32	$6{,}30\cdot10^{-1}$	$3{,}38\cdot10^{3}$	$3{,}55\cdot10^{2}$	$5{,}63\cdot10^{-1}$	$5{,}63\cdot10^{-1}$
40	$5{,}09\cdot10^{-1}$	$3{,}07\cdot10^{7}$	$2{,}44\cdot10^{6}$	$5{,}03\cdot10^{-1}$	$4{,}85\cdot10^{-1}$
48	$4{,}04\cdot10^{-1}$	parep	$3{,}46\cdot10^{10}$	$2{,}47\cdot10^{6}$	$3{,}64\cdot10^{-1}$
56	$3{,}17\cdot10^{-1}$	—	$1{,}02\cdot10^{15}$	$2{,}29\cdot10^{2}$	$3{,}10\cdot10^{-1}$
64	$2{,}47\cdot10^{-1}$	—	parep	$1{,}87\cdot10^{4}$	$2{,}23\cdot10^{-1}$
72	$1{,}92\cdot10^{-1}$	—	—	$1{,}73\cdot10^{6}$	$1{,}72\cdot10^{-1}$
80	$1{,}49\cdot10^{-1}$	—	—	parep	$1{,}44\cdot10^{-1}$
...	...	...	...	...	...
256	$4{,}97\cdot10^{-4}$	—	—	—	$4{,}80\cdot10^{-4}$
...	...	...	...	...	...
512	$1{,}23\cdot10^{-7}$	—	—	—	$1{,}15\cdot10^{-7}$

a otra $S_h = E - \tau_h C$ es mayor que la unidad para algunos valores de k.

En efecto, como S es un operador autoconjugado en H, entonces $\| S_k \| = \sup_{\|x\|=1} |(S_k x, x)|$. Utilizando las cotas γ_1 y γ_2 del operador C

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \gamma_1 > 0,$$

hallaremos

$$(1 - \tau_h\gamma_2)\, E \leqslant E - \tau_h C \leqslant (1 - \tau_h\gamma_1)\, E.$$

Sustituyamos aquí τ_h de (15) y tengamos en consideración las igualdades $1 - \rho_0 = \tau_0\gamma_1$, $1 + \rho_0 = \tau_0\gamma_2$. Obtendremos

$$-\frac{\rho_0(1-\mu_k)}{1+\rho_0\mu_k} E \leqslant S_k \leqslant \frac{\rho_0(1+\mu_k)}{1+\rho_0\mu_k} E$$

y, por consiguiente,

$$\| S_k \| = \begin{cases} \dfrac{\rho_0(1+\mu_k)}{1+\rho_0\mu_k} < 1, & \mu_k \geqslant 0, \\ \dfrac{\rho_0(1-\mu_k)}{1+\rho_0\mu_k}, & \mu_k < 0. \end{cases}$$

De aquí se deduce, que $\| S_k \| > 1$ para $\mu_k < -(1 - \rho_0)/(2\rho_0)$. Ya que $\mu_k \in \mathfrak{M}_n$, entonces

$$-\cos\frac{\pi}{2n} \leqslant \mu_k \leqslant -\cos\frac{2n-1}{2n}\pi = \cos\frac{\pi}{2n}, \quad k = 1, 2, \ldots, n,$$

y, por lo tanto, para una cantidad de números k grande la norma $\| S_k \| > 1$ (la cantidad de estos números k es aproximadamente igual a $n/2$). Por eso, si se utilizan uno tras otro gran cantidad de parámetros τ_k, para los cuales la norma del operador S_k es mayor que la unidad, entonces puede ocurrir una acumulación de error computacional y el crecimiento de las aproximaciones iterativas, lo cual sirve como causa de inestabilidad computacional del método.

El teorema 1 expresa realmente la estabilidad del esquema iterativo según los datos iniciales. En el caso de un proceso computacional real es necesario investigar también la estabilidad del esquema iterativo respecto del segundo miembro, por cuanto los errores de redondeo se pueden interpretar como una perturbación del segundo miembro del esquema iterativo en cada iteración.

Si tomamos en consideración el error de redondeo, entonces en lugar de la ecuación homogénea para el error equivalente x_k obtendremos una ecuación no homogénea

$$x_{k+1} = S_{k+1} x_k + \tau_{k+1} \varphi_{k+1}, \quad k = 0, 1, \ldots \qquad (16)$$

Aquí $x_k = D^{1/2} (\bar{y}_k - u)$, donde $\bar{y}_k$ es la aproximación iterativa real.

Resolviendo la ecuación (16) hallamos $x_n = T_{n,0} x_0 + \sum_{j=1}^{n} \tau_j T_{n,j} \varphi_j$, donde $T_{n,j} = \prod_{i=j+1}^{n} S_i$, $T_{n,n} = E$.

De aquí obtenemos la siguiente estimación:

$$\| x_n \| \leqslant \| T_{n,0} \| \| x_0 \| + \sum_{j=1}^{n} \tau_j \| T_{n,j} \| \max_{1 \leqslant j \leqslant n} \| \varphi_j \|. \qquad (17)$$

La estimación de la norma del operador $T_{n,0}$ no dependen del ordenamiento del conjunto $\mathfrak{M}_n$ y para cualquier sucesión de los parámetros de Chebishev τ_k tenemos $\| T_{n,0} \| \leqslant q_n$. La estimación para $\sum_{j=1}^{n} \tau_j \| T_{n,j} \|$ depende del ordenamiento del conjunto $\mathfrak{M}_n$. De (17) se deduce, que el conjunto $\mathfrak{M}_n$ debe ser ordenado de manera tal que la suma indicada toma el valor mínimo.

El siguiente lema indica el valor mínimo posible de esta suma.

LEMA 1. *Si* γ_1 y γ_2 *son las cotas exactas del operador* C, *entonces para cualquier ordenamiento del conjunto* $\mathfrak{M}_n$ *tiene lugar la estimación*

$$\sum_{i=1}^{n} \tau_j \| T_{n,j} \| \geqslant \frac{1-q_n}{\gamma_1}.$$

Realmente, de la definición de los operadores $T_{n,j}$ obtenemos

$$\tau_j T_{n,j} = (T_{n,j} - T_{n,j-1}) C^{-1}, \quad \sum_{j=1}^{n} \tau_j T_{n,j} = (E - T_{n,0}) C^{-1}.$$

Puesto que

$$\| (E - T_{n,0}) C^{-1} \| = \Big\| \sum_{j=1}^{n} \tau_j T_{n,j} \Big\| \leqslant \sum_{j=1}^{n} \tau_j \| T_{n,j} \|,$$

entonces es suficiente estimar la norma del operador $(E - T_{n,0}) C^{-1}$. Este operador es autoconjugado en H, y si γ_1 y γ_2 son las cotas del operador C, entonces

$$\| (E - T_{n,0}) C^{-1} \| \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \left| \frac{1 - q_n T_n \left(\frac{1 - \tau_0 t}{\rho_0} \right)}{t} \right| =$$

$$= \frac{1 - q_n T_n \left(\frac{1 - \tau_0 \gamma_1}{\rho_0} \right)}{\gamma_1} = \frac{1 - q_n}{\gamma_1}.$$

De esta forma, está mostrado que para cualquier $x \in H$ tiene lugar la estimación

$$\| (E - T_{n,0}) C^{-1} x \| \leqslant \frac{1 - q_n}{\gamma_1} \| x \|. \tag{18}$$

Como γ_1 es la cota exacta del operador autoconjugado C, entonces γ_1 coincide con el valor propio mínimo del operador C. Sustituyendo en (18) la función propia, correspondiente al valos propio mínimo del operador C, en lugar de x, se obtiene que en (18) se alcanza la igualdad. Por lo tanto, hemos obtenido la estimación $\| (E - T_{n,0}) C^{-1} \| = (1 - q_n)/\gamma_1$. El lema está demostrado.

5. Construcción de una sucesión óptima de parámetros iterativos*).

*) El método de ordenamiento de los parámetros esta dado según los trabajos: véase E. S. Nikolaev, A. A. Samarski (Ch V M y MF, 12 No. 4, 1972) para cualquier n y [8] para $n = 2^P$.

5.1 CASO $n=2^P$. El orden de utilización de los parámetros iterativos τ_k en el método de Chebishev influye esencialmente en la convergencia del método. Por eso aparece el problema de construir la mejor sucesión de parámetros iterativos, la cual garantice la influencia mínima del error computacional del método. Como la sucesión de los parámetros se determina por un ordenamiento del conjunto $\mathfrak{M}_n$, entonces es necesario construir el ordenamiento óptimo de este conjunto $\mathfrak{M}_n$.

Exponemos la resolución de este problema. Supongamos primeramente que el número de iteraciones es una potencia de 2: $n = 2^P$. Designemos mediante θ_m un conjunto que consiste de m números enteros:

$$\theta_m = \{\theta_1^{(m)}, \theta_2^{(m)}, \ldots, \theta_m^{(m)}\}.$$

Partiendo del conjunto $\theta_1 = \{1\}$, construiremos el conjunto θ_{2^P} por la regla siguiente. Supongamos que el conjunto θ_m está construido. Entonces hallamos el conjunto θ_{2m} mediante las fórmulas

$$\theta_{2m} = \{\theta_{2i}^{(2m)} = 4m - \theta_i^{(m)},\ \theta_{2i-1}^{(2m)} = \theta_i^{(m)},$$
$$i = 1, 2, \ldots, m\},\quad m = 1, 2, 4, \ldots, 2^{P-1}. \tag{19}$$

No es difícil cerciorarse de que el conjunto θ_{2^k} consiste de los números impares desde 1 hasta $2^{k+1} - 1$.

Utilizando la construcción del conjunto θ_{2^P}, ordenaremos el conjunto $\mathfrak{M}_{2^P}$ de la siguiente forma:

$$\mathfrak{M}_n^* = \left\{-\cos\beta_i,\ \beta_i = \frac{\pi}{2n}\theta_i^{(n)},\ i = 1, 2, \ldots, n\right\},\ n = 2^P. \tag{20}$$

Este es el ordenamiento buscado del conjunto $\mathfrak{M}_n$ en el caso, cuando $n = 2^P$. Para la sucesión de parámetros iterativos correspondiente a este ordenamiento está demostrado la estimación

$$\sum_{j=1}^{n} \tau_j \| T_{n,j} \| \leqslant \frac{1 - q_n}{\gamma_1}.$$

Comparando esta estimación con la del lema 1, nos cercioramos de que el ordenamiento construido del conjunto $\mathfrak{M}_n^{(*)}$ realmente proporciona una influencia mínima del error computacional en la convergencia del método de Chebishev.

Citemos algunos ejemplos de construcción del conjunto θ_n.

1) $n = 8$

$$\theta_1 = \{1\}, \quad \theta_2 = \{1, 3\}, \quad \theta_4 = \{1, 7, 3, 5\},$$
$$\theta_8 = \{1, 15, 7, 9, 3, 13, 5, 11\}.$$

El conjunto θ_8 está construído. Por la fórmula (20) se ordena el conjunto $\mathfrak{M}_8^{(*)}$.

2) $n = 16$

Aplicando el conjunto θ_8 hallado más arriba, construimos el conjunto θ_{16} por las fórmulas (19):

$\theta_{16} = \{1, 31, 15, 17, 7, 25, 9, 23, 3, 29, 13, 19, 5, 27, 11, 21\}$.

3) $n = 32$.

$$\theta_{32} = \{1, 63, 31, 33, 15, 49, 17, 47, 7, 57, 25, 39, 9, 55,$$
$$23, 41, 3, 61, 29, 35, 13, 51, 19, 45, 5, 59, 27, 37,$$
$$11, 53, 21, 43\}.$$

De las fórmulas (19) se deduce una regla simple de transición del conjunto θ_m al conjunto θ_{2m}: $\theta_{2i-1}^{(2m)} = \theta_i^{(m)}$ y la suma de dos números adyacentes es igual a $4m$:

$$\theta_{2i-1}^{(2m)} + \theta_{2i}^{(2m)} = 4m, \quad i = 1, 2, \ldots, m.$$

Una regla análoga de transición se aplica en el caso general, a cuyo examen nosotros pasamos.

5.2 CASO GENERAL. Supongamos que el número n de iteraciones es un número entero cualquiera. Describamos el proceso de construcción del conjunto θ_n. Las etapas elementales de este proceso son los pasos del conjunto θ_m al conjunto θ_{2m} y del conjunto θ_{2m} al conjunto θ_{2m+1}, donde m es un número entero arbitrario.

Formulemos las reglas de transición de un conjunto a otro.

1) El paso de θ_{2m} a θ_{2m+1} consiste en la adición del número impar $2m + 1$ a los elementos del conjunto θ_{2m}.

2) La transición de θ_m a θ_{2m} se realiza de la siguiente forma. Si después de esta transición sigue la de θ_{2m} a θ_{4m} o si la transición de θ_m a θ_{2m} es el último paso en el proceso de construcción del conjunto θ_n, entonces se utilizan las fórmulas citadas más arriba:

$$\theta_{2i-1}^{(2m)} = \theta_i^{(m)}, \ \theta_{2i-1}^{(2m)} + \theta_{2i}^{(2m)} = 4m, \ i = 1, 2, \ldots, m. \quad (21)$$

Si después de la transición de θ_m a θ_{2m} sigue la transición de θ_{2m} a θ_{2m+1}, entonces se utilizan las fórmulas

$$\theta_{2i-1}^{(2m)} = \theta_i^{(m)}, \quad \theta_{2i-1}^{(2m)} + \theta_{2i}^{(2m)} = 4m + 2, \quad i = 1, 2, \ldots . m. \qquad (22)$$

Empleando estas reglas y alternando debidamente las transiciones de un conjunto con un número par de elementos a un conjunto con un número impar de elementos y de un conjunto de m elementos a un conjunto de $2m$ elementos, se puede construir el conjunto θ_n para cualquier n, partiendo de $\theta_1 = \{1\}$.

Mostremos algunos ejemplos:

1) $n = 15$. En este caso el paso de θ_1 a θ_n se produce según la cadena siguiente:

$$\theta_1 \to \theta_2 \to \theta_3 \to \theta_6 \to \theta_7 \to \theta_{14} \to \theta_{15}.$$

De acuerdo con las reglas expuestas, las transiciones de θ_1 a θ_2, de θ_3 a θ_6 y de θ_7 a θ_{14} se realizan por las fórmulas (22), y durante la transición de θ_2 a θ_3, de θ_6 a θ_7 y de θ_{14} a θ_{15} hay que añadir el correspondiente número impar al conjunto inicial. Esto da:

$\theta_1 = \{1\}$, $\theta_2 = \{1, 5\}$, $\theta_3 = \{1, 5, 3\}$,

$\theta_6 \{1, 13, 5, 9, 3, 11\}$, $\theta_7 = \{1, 13, 5, 9, 3, 11, 7\}$,

$\theta_{14} = \{1, 29, 13, 17, 5, 25, 9, 21, 3, 27, 11, 19, 7, 23\}$,

$\theta_{15} = \{1, 29, 13, 17, 5, 25, 9, 21, 3, 27, 11, 19, 7, 23, 15\}$.

El conjunto $\mathfrak{M}_{15}^{(*)}$ se ordena según la fórmula (20).

2) $n = 25$. A este caso corresponde la cadena

$$\theta_1 \to \theta_2 \to \theta_3 \to \theta_6 \to \theta_{12} \to \theta_{24} \to \theta_{25}.$$

y las transiciones de θ_1 a θ_2 y de θ_{12} a θ_{24} se ejecutan por las fórmulas (22), las transiciones de θ_3 a θ_6 y de θ_6 a θ_{12} por las fórmulas (21) y las de θ_2 a θ_3 y de θ_{24} a θ_{25} se realizan por la adición de un número impar. Obtendremos

$\theta_1 = \{1\}$, $\theta_2 = \{1, 5\}$, $\theta_3 = \{1, 5, 3\}$, $\theta_6 = \{1, 11, 5, 7, 3, 9\}$,

$\theta_{12} = \{1, 23, 11, 13, 5, 19, 7, 17, 3, 21, 9, 15\}$,

$\theta_{24} = \{1, 49, 23, 27, 11, 39, 13, 37, 5, 45, 19, 31, 7, 43, 17, 33, 3, 47, 21, 29, 9, 41, 15, 35\}$,

$\theta_{25} = \{1, 49, 23, 27, 11, 39, 13, 37, 5, 45, 19, 31, 7, 43, 17, 33, 3, 47, 21, 29, 9, 41, 15, 35, 25\}$.

El procedimiento de construcción del conjunto θ_n expuesto más arriba puede ser formalizado para n arbitrario. Para esto representemos n en forma de un desarrollo según las potencias de 2 con exponentes enteros k_j:

$$n = 2^{k_1} + 2^{k_2} + \ldots + 2^{k_s}, \quad k_j \leqslant k_{j-1} - 1,$$

$$j = 2, 3, \ldots, s.$$

Formemos las siguientes cantidades:

$$n_j = \sum_{i=1}^{j} 2^{k_i - k_j}, \quad j = 1, 2, \ldots, s,$$

y pongamos $n_{s+1} = 2n + 1$. Por las fórmulas (23) construyamos el conjunto θ_{n_j}:

$$\theta_{n_j} = \{\theta_i^{(n_j)} = \theta_i^{(n_j-1)}, \ \theta_{n_j}^{(n_j)} = n_j, \ i = 1, 2. \ldots, n_j - 1\}, \quad (23)$$

para $j = 1$ eligimos $\theta_1 = \{1\}$. Después por la fórmula (24) se construyen los conjuntos

$$\theta_{2m} = \{\theta_{2i}^{(2m)} = 4m - \theta_i^{(m)}, \ \theta_{2i-1}^{(2m)} = \theta_i^{(m)}, \ i = 1, 2, \ldots, m\} \quad (24)$$

para $m = n_j, 2n_j, 4n_j, \ldots, [(n_{j+1} - 1)/4]$, donde $[a]$ es la parte entera de a. Si $[(n_{j+1} - 1)/4] < n_j$, entonces los cálculos para la fórmula (24) no se realizan y se pasa a la siguiente etapa. Si $j = s$, entonces el conjunto θ_n necesario ya está construido. De otro modo suponemos que $m = (n_{j+1} - 1)/2$ y construiemos el conjunto

$$\theta_{2m} = \{\theta_{2i}^{(2m)} = 4m + 2 - \theta_i^{(m)}, \ \theta_{2i-1}^{(2m)} = \theta_i^{(m)}, \ i = 1, 2, \ldots, m\}. \quad (25)$$

Luego j aumenta en la unidad y el proceso se repite partiendo de la fórmula (23). Debido será construido el conjunto θ_n. El conjunto $\mathfrak{M}_n^*$ se ordena de acuerdo con la fórmula (20).

Para el caso $n = 2^P$ el algoritmo (23)-(25) se simplifica y pasa a ser el algoritmo descrito por la fórmula (19). En realidad, para $n = 2^P$ obtenemos $s = 1$, $k_1 = P$, $n_1 = 1$, $n_{s+1} = 2^{P+1} - 1$. Por lo tanto, en el algoritmo (23)-(25) j toma un único valor, igual a la unidad, y los cálculos se efectúa mediante la fórmula (24) para $m = 1, 2, 4, \ldots \ldots, 2^{P-1}$.

Ilustremos aquí la calidad del ordenamiento construido del conjunto: $\mathfrak{M}_n^{(*)}$ en el ejemplo examinado en el punto 4 del § 2. El número prefijado de iteraciones n se ha cambiado

desde 16 hasta 512 con el paso 8. Para cada n la exactitud real alcanzada después de ser efectuadas n iteraciones, no ha superado la exactitud teórica q_n ($q_{512} = 1{,}23 \cdot 10^{-7}$), y el proceso ha sido «sin paradas de reparación» (véase la tabla 5).

§ 3. Método de iteración simple

1. Elección del parámetro iterativo. En el § 2 fue resuelto el problema sobre la construcción del conjunto óptimo de los parámetros iterativos τ_k para el esquema de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f,\ k=0,\ 1,\ \ldots,\ y_0 \in H$$

suponiendo que el operador $DB^{-1}A$ es autoconjugado en H y siendo dadas las constantes γ_1 y γ_2 de equivalencia energética de los operadores D y $DB^{-1}A$:

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D,\quad \gamma_1 > 0. \tag{1}$$

Obtendremos ahora la solución de este problema bajo la restricción complementaria $\tau_k \equiv \tau$, es decir suponiendo, que los parámetros iterativos τ_k no dependen del número k de iteraciones. Este problema aparece durante la búsqueda del parámetro iterativo τ para el esquema estacionario de dos capas.

$$B\frac{y_{k+1}-y_k}{\tau} + Ay_k = f,\ k=0,\ 1,\ \ldots. \tag{2}$$

Recordemos la formulación del problema indicado más arriba: entre los polinomios de grado n del tipo $Q_n(t) = \prod_{j=1}^{n}(1-\tau_j t)$ hallar el polinomio con menor desviación respecto del cero en el segmento $[\gamma_1, \gamma_2]$. En virtud de la restricción hecha el polinomio $P_n(t)$ tiene la forma

$$P_n(t) = (1-\tau t)^n.$$

Por eso el problema planteado más arriba es equivalente al siguiente: entre los polinomios de primer grado, que toman el valor uno en el punto $t = 0$, hallar el polinomio con menor desviación del cero en el intervalo $[\gamma_1, \gamma_2]$.

Este problema es un caso particular del problema examinado en el § 2. En el caso dado $n = 1$ y de los resultados del

punto 1 del § 2 se deduce, que el polinomio buscado tiene la forma

$$Q_1(t) = q_1 T_1\left(\frac{1-\tau_0 t}{\rho_0}\right), \quad \tau_0 = \frac{2}{\gamma_1+\gamma_2},$$

$$\rho_0 = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

donde

$$q_1 = \frac{2\rho_1}{1+\rho_1^2} = \rho_0, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}.$$

Aquí $T_1(x)$ es el polinomio de Chebishev de primer género. Ya que $T_1(x) = x$, el polinomio $Q_1(t)$ tiene la forma

$$Q_1(t) = 1 - \tau_0 t, \quad \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |Q_1(t)| = q_1 = \rho_0,$$

por eso

$$P_n(t) = (1 - \tau_0 t)^n.$$

De esta forma, hemos hallado el valor óptimo del parámetro τ para el esquema (2):

$$\tau = \tau_0 = 2/(\gamma_1 + \gamma_2). \tag{3}$$

Como la norma del operador de resolución solvente $T_{n,0}$ para el esquema (2) (véase el punto 3 del § 1) se estima de la siguiente forma:

$$\| T_{n,0} \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |P_n(t)|,$$

entonces para $\tau = \tau_0$ obtenemos la estimación $\| T_{n,0} \| \leqslant \rho_0^n$. De aquí se deduce una estimación para el error z_n en H_D:

$$\| z_n \|_D \leqslant \rho_0^n \| z_0 \|_D. \tag{4}$$

El método iterativo (2), (3) se llama *método de iteración simple*.

Así, hemos demostrado

TEOREMA 2. *Supongamos que el operador autoconjugado* $DB^{-1}A$ *satisface las condiciones* (1). *El método de iteración simple* (2), (3) *converge en* H_D *y para el error tiene lugar la estimación* (4). *Para el número de iteraciones es válida la estimación* $n \geqslant n_0(\varepsilon)$, *donde* $n_0(\varepsilon) = \ln \varepsilon / \ln \rho_0$.

OBSERVACIÓN. Al igual que para el método de Chebishev, la estimación a priori del error para el método de iteración

simple es inmejorable en el caso de un espacio de dimensión finita.

Comparemos el número de iteraciones para el método de Chebishev y el método de iteración simple. Del teorema 1 en el caso de ξ pequeños tenemos la siguiente estimación para el número de iteraciones del método de Chebishev:

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \frac{\ln 0{,}5\varepsilon}{\ln \rho_1} \approx \frac{1}{2\sqrt{\xi}} \ln \frac{2}{\varepsilon}.$$

Del teorema 2 obtenemos una estimación para el número de iteraciones del método de iteración simple

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \frac{\ln \varepsilon}{\ln \rho_0} \approx \frac{1}{2\xi} \ln \frac{1}{\varepsilon}.$$

De estas estimaciones se deduce, que para $\xi \ll 1$ el número de iteraciones del método de Chebishev es esencialmente menor que el número de iteraciones del método de iteración simple. Por ejemplo, para $\xi = 0{,}01$ el número de iteraciones para el método de iteración simple es aproximadamente 10 veces mayor que para el método de Chebishev.

2. Estimación de la norma del operador de transición. En el punto 1 del § 3 fue investigada la velocidad de convergencia del método de iteración simple. Con esto el método de iteración simple se examinó como un caso particular del método de Chebishev. De las consideraciones metodológicas será útil estudiar la convergencia del método de iteración simple independientemente del método de Chebishev.

Así pues, supongamos que para encontrar una solución aproximada de la ecuación

$$Au = f$$

se utiliza el esquema de dos capas (2)

$$B\frac{y_{k+1}-y_k}{\tau} + Ay_k = f, \quad k = 0, 1, \ldots, \quad y_0 \in H. \qquad (5)$$

Para estudiar la convergencia del esquema (5) pasaremos al problema para el error equivalente $x_k = D^{1/2}z_k$:

$$x_{k+1} = Sx_k, \quad k = 0, 1, \ldots, S = E - \tau C, \qquad (6)$$

donde $C = D^{1/2}B^{-1}AD^{-1/2}$. Utilizando (6) hallamos la expresión explícita para x_n mediante x_0: $x_n = S^n x_0$, de

la cual se deduce una estimación para la norma del error z_n en H_D

$$\| z_n \|_D = \| x_n \| \leqslant \| S^n \| \| x_0 \| = \| S^n \| \| z_0 \|_D. \tag{7}$$

Supondremos, que el operador $DB^{-1}A$ es autoconjugado en H y que están prefijadas las constantes γ_1 y γ_2 en las desigualdades (1). Para estas suposiciones el operador C, y junto con él el operador S, son autoconjugados en H y γ_1, γ_2 son las cotas del operador C:

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \gamma_1 > 0, \quad C = C^*. \tag{8}$$

En virtud de la autoconjugación del operador S tiene lugar la igualdad $\| S^n \| = \| S \|^n$. Por eso de la estimación (7) se deduce, que es necesario elegir el parámetro iterativo τ de la condición de mínimo según τ de la norma del operador de transición $S = E - \tau C$.

Tiene lugar

LEMA 2. *Sea* $S = E - \tau C$ *y supongamos que se cumplen las condiciones* (8). *La norma del operador* S *es mínima para* $\tau = \tau_0 = 2/(\gamma_1 + \gamma_2)$, *y tiene lugar la estimación*

$$\| S \| = \| E - \tau_0 C \| = \rho_0, \quad \rho_0 = (1 - \xi)/(1 + \xi),$$
$$\xi = \gamma_1/\gamma_2.$$

Realmente, puesto que S es un operador autoconjugado en H, entonces de la definición de la norma obtenemos

$$\| S \| = \sup_{x \neq 0} \frac{|(Sx, x)|}{(x, x)} = \sup_{x \neq 0} \left| 1 - \tau \frac{(Cx, x)}{(x, x)} \right| = \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |1 - \tau t|.$$

Ya que $\varphi(t) = 1 - \tau t$ es una función lineal, entonces el valor máximo en módulo de $\varphi(t)$ en el segmento $[\gamma_1, \gamma_2]$ se puede alcanzar solamente en los extremos del segmento. Los cálculos directos proporcionan

$$\| S \| = \max(|1 - \tau\gamma_1|, |1 - \tau\gamma_2|) =$$
$$= \begin{cases} \varphi_1(\tau) = 1 - \tau\gamma_1, & 0 \leqslant \tau \leqslant \tau_0, \\ \varphi_2(\tau) = \tau\gamma_2 - 1, & \tau_0 \leqslant \tau, \end{cases}$$

donde τ_0 está indicado en el lema. Como la función $\varphi_1(\tau)$ decrece sobre el segmento $[0, \tau_0]$, y $\varphi_2(\tau)$ crece para $\tau \geqslant \tau_0$, entonces el mínimo de la norma del operador S se alcanza para $\tau = \tau_0$ y es igual a $\rho_0 = 1 - \tau_0\gamma_1 = \tau_0\gamma_2 - 1 = (1 - \xi)/(1 + \xi)$, $\xi = \gamma_1/\gamma_2$. El lema está demostrado.

Del lema 2 y de la estimación (7) se deduce que cuando $\tau = \tau_0$ para el error del esquema iterativo (5) es cierto la estimación

$$\| z_n \|_D \leqslant \rho_0^n \| z_0 \|_D.$$

De esta forma, hemos obtenido otra demostración del teorema 2 formulado más arriba sobre la convergencia del método de iteración simple. Los ejemplos de elección del operador D, para el cual se cumple la condición de autoconjugación del operador $DB^{-1}A$, son examinados en el punto 3 del § 2.

§ 4. Caso no autoconjugado. Método de iteración simple

1. Planteamiento del problema. En los §§ 2, 3 fueron construidos métodos iterativos de dos capas para la resolución aproximada de la ecuación operacional lineal

$$Au = f \tag{1}$$

con un operador no degenerado A, definido en el espacio de Hilbert real H. Se supuso que los operadores A, B y D son tales, que el operador $DB^{-1}A$ era autoconjugado en H y que eran prefijadas las constantes γ_1 y γ_2 de equivalencia energética de los operadores D y $DB^{-1}A$, además $\gamma_1 > 0$.

Con estas suposiciones fue resuelto el problema de elección óptima de los parámetros iterativos y fueron construidos el método de Chebishev y el método de iteración simple. En el punto 3 del § 2 fueron examinados algunos ejemplos de elección del operador D y halladas las condiciones de autoconjugación del operador $DB^{-1}A$ para cada elección concreta del operador D.

Es evidente, que si son prefijados los operadores A y B, no siempre se puede indicar un operador D, para el cual el operador $DB^{-1}A$ sea autoconjugado en H. Por consiguiente, es necesario estudiar los métodos iterativos también en el caso no autoconjugado.

En este parrafo se estudia el método de iteración simple para el caso no autoconjugado. Serán examinados algunos procedimientos para elegir el parámetro iterativo en función del volúmen de información a priori acerca de los operadores del esquema iterativo.

Así pues, sea el operador $DB^{-1}A$ *no autoconjugado* en H. Para la resolución aproximada de la ecuación (1) examinemos el esquema iterativo implícito de dos capas

$$B\frac{y_{k+1}-y_k}{\tau}+Ay_k=f,\ k=0,\ 1,\ \ldots,\quad y_0\in H. \tag{2}$$

Para *investigar* la convergencia del esquema (2), como es *frecuente*, *pasaremos al problema respecto al error equivalente* $x_k=D^{1/2}z_k$

$$x_{k+1}=Sx_k,\quad k=0,\ 1,\ \ldots,\quad S=E-\tau C, \tag{3}$$

donde $C=D^{1/2}B^{-1}AD^{-1/2}$. En virtud de las suposiciones hechas más arriba el operador C no es autoconjugado en H. Del cambio hecho y de la ecuación (3) obtenemos

$$x_n=S^n\ x_0,\quad \|x_n\|=\|z_n\|_D\leqslant\|S^n\|\,\|x_0\|= \\ =\|S^n\|\,\|z_0\|_D. \tag{4}$$

Por consiguiente, el parámetro iterativo τ debe ser elegido de la condición de mínimo según τ de la norma del operador de resolución S^n.

2. Minimización de la norma del operador de transición.

2.1 PRIMER CASO. Obtengamos una estimación para la norma del operador S^n. Ya que para cualquier operador tiene lugar la estimación $\|S^n\|\leqslant\|S\|^n$, entonces el primer procedimiento de elección del parámetro τ consiste en encontrarlo de la condición de mínimo de la norma del operador de transición S. Obtendremos dos tipos de estimaciones de la norma del operador S en dependencia del volúmen de información a priori respecto al operador C.

En el primer caso se supone, que la información a priori consiste en prefijar las constantes γ_1 y γ_2 de las desigualdades

$$\gamma_1\ (x,\ x)\leqslant(Cx,\ x),\quad (Cx,\ Cx)\leqslant\gamma_2\ (Cx,\ x), \\ \gamma_1>0. \tag{5}$$

Si $C=C^*$, entonces γ_1 y γ_2 son las cotas del operador C.

LEMA 3. *Sean dadas* γ_1 y γ_2 *en las desigualdades* (5), *entonces para la norma del operador* $S=E-\tau C$ *con* $\tau=1/\gamma_2$ *es válida la estimación*

$$\|S\|\leqslant\rho,\ \rho=\sqrt{1-\xi},\ \xi=\gamma_1/\gamma_2.$$

En efecto, utilizando (5), obtenemos

$$\|Sx\|^2=\|x-\tau Cx\|^2=(x,\ x)-2\tau\ (Cx,\ x)+ \\ +\tau^2\ (Cx,\ Cx)\leqslant(x,\ x)-2\tau\ (Cx,\ x)+\tau^2\gamma_2\ (Cx,\ x)= \\ =\|x\|^2-\tau\ (2-\tau\gamma_2)\ (Cx,\ x).$$

De aquí se deduce, que si se cumple la condición $\tau(2 - \tau\gamma_2) > 0$, es decir $0 < \tau < 2/\gamma_2$, entonces la norma del operador S será menor que la unidad. Supongamos que esta condición se cumple, entonces, empleando (5), obtendremos

$$\| Sx \|^2 \leqslant [1 - \tau\gamma_1 (2 - \tau\gamma_2)] \| x \|^2$$

y, por lo tanto,

$$\| S \|^2 = \sup_{x \neq 0} \frac{\| Sx \|^2}{\| x \|^2} \leqslant 1 - \tau\gamma_1 (2 - \tau\gamma_2).$$

La función $\varphi(\tau) = 1 - \tau\gamma_1 (2 - \tau\gamma_2)$ tiene el mínimo en el punto $\tau = 1/\gamma_2$, igual a $\varphi(1/\gamma_2) = 1 - \xi$, donde $\xi = \gamma_1/\gamma_2$. Por eso para el valor indicado del parámetro τ para la norma del operador S es válida la estimación $\| S \| \leqslant \sqrt{1 - \xi}$. El lema está demostrado.

Sustituyendo el operador $C = D^{-1/2} (DB^{-1}A) D^{-1/2}$ en (5), obtendremos, que las desigualdades (5) son equivalentes a las desigualdades siguientes:

$$\gamma_1 (Dx, x) \leqslant (DB^{-1}Ax, x),$$
$$(DB^{-1}Ax, B^{-1}Ax) \leqslant \gamma_2 (DB^{-1}Ax, x), \quad \gamma_1 > 0. \qquad (6)$$

Sustituyendo en (4) la estimación para la norma del operador S, obtenido en el lema 3, hallamos

$$\| z_n \|_D \leqslant \rho^n \| z_0 \|_D, \quad \rho = \sqrt{1 - \xi}. \qquad (7)$$

TEOREMA 3. *Sean* γ_1 y γ_2 *las constantes de las desigualdades* (6). *Para el valor del parámetro iterativo* $\tau = 1/\gamma_2$ *el método de iteración simple* (2) *converge en* H_D, *y para el error* z_n *tiene lugar la estimación* (7). *Para el número de iteraciones es cierta la estimación* $n \geqslant n_0(\varepsilon)$, *donde* $n_0(\varepsilon) = \ln \varepsilon / \ln \rho$, $\rho = \sqrt{1 - \xi}$, $\xi = \gamma_1/\gamma_2$.

Citemos ejemplos de la elección del operador D y el tipo concreto de las desigualdades (6). En la tabla 6 se exponen las suposiciones respecto a los operadores A y B, y se indican el operador D y el tipo de las desigualdades (6). Para obtener el tipo concreto de las desigualdades (6) nosotros partimos tanto de las mismas desigualdades (6) como de las desigualdades equivalentes

$$\gamma_1 (DA^{-1}Bx, A^{-1}Bx) \leqslant (DA^{-1}Bx, x), \quad (Dx, x) \leqslant$$
$$\leqslant \gamma_2 (DA^{-1}Bx, x), \qquad (8)$$

obtenidas de (6) valiéndose de la sustitución $x = A^{-1}By$.

Señalemos las desigualdades

$$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx), \quad (Ax, Ax) \leqslant \gamma_2 (Ax, Bx), \quad \gamma_1 > 0.$$

Si estas condiciones son cumplidas, entonces para los casos particulares examinados en la tabla 6, se puede tomar

Tabla 6

A y B	D	Desigualdades
1) $A = A^* > 0$, $B \neq B^* > 0$	$A \circ B^* A^{-1} B$ $A^2 \circ B^* B$	$\gamma_1 (Bx, A^{-1} Bx) \leqslant (Bx, x)$, $(Ax, x) \leqslant \gamma_2 (Bx, x)$ $\gamma_1 (Bx, Bx) \leqslant (Ax, Bx)$, $(Ax, Ax) \leqslant \gamma_2 (Ax, Bx)$
2) $A \neq A^* > 0$, $B = B^* > 0$	$B \circ A^* B^{-1} A$ $B^2 \circ A^* A$	$\gamma_1 (Bx, x) \leqslant (Ax, x)$, $(Ax, B^{-1} Ax) \leqslant \gamma_2 (Ax, x)$ $\gamma_1 (Bx, Bx) \leqslant (Ax, Bx)$, $(Ax, Ax) \leqslant \gamma_2 (Ax, Bx)$
3) $A \neq A^* > 0$, $B \neq B^* > 0$	$A^* A \circ B^* B$	$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx)$, $(Ax, Ax) \leqslant \gamma_2 (Ax, Bx)$
4) $A = A^*$, $B = B^*$ $AB \neq BA$	$A^2 \circ B^2$	$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx)$, $(Ax, Ax) \leqslant \gamma_2 (Ax, Bx)$

en calidad de operador D el operador A^2, si $A = A^*$, o el operador A^*A. Para tal elección del operador D la norma del error z_n en H_D puede ser calculada en el proceso de las iteraciones

$$\| z_n \|_D^2 = (Dz_n, z_n) = (Az_n, Az_n) = \| r_n \|^2, \quad r_n = Ay_n - f.$$

Volvamos a la estimación de la norma del operador S. Si el operador C es autoconjugado en H, entonces en virtud de (5) él es definido positivo, y por lo tanto, existe la raíz cuadrada del operador C.

Poniendo $x = C^{-1/2} y$ en la segunda de las desigualdades (5), obtenemos, que las desigualdades (5) son equivalentes a las desigualdades

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \gamma_1 > 0.$$

De lema 2 para estas suposiciones se infiere la siguiente estimación de la norma del operador S: $\| S \| \leqslant \rho_0$, $\rho_0 = (1 - \xi)/(1 + \xi)$, $\xi = \gamma_1/\gamma_2$.

Comparando esta estimación con la obtenida en el lema 3, nos cercioramos de que la estimación del lema 3 es burda y no pasa a ser la estimación del lema 2, cuando el operador C es autoconjugado en H.

2.2 SEGUNDO CASO. Obtendremos ahora otra estimación para la norma del operador de transición S, la cual se convierte en la estimación del lema 2, cuando C es un operador autoconjugado en H. Para esto aumentemos el volúmen de información a priori respecto al operador C, suponiendo que están prefijados los tres números γ_1, γ_2 y γ_3:

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \| C_1 \| \leqslant \gamma_3, \quad \gamma_1 > 0, \quad \gamma_3 \geqslant 0, \tag{9}$$

donde $C_1 = 0{,}5\,(C - C^*)$ es la parte no autoconjugada del operador C.

Tiene lugar

LEMA 4. *Sean prefijados* γ_1, γ_2 y γ_3 *en las desigualdades* (9). *Entonces para la norma del operador* $S = E - \tau C$ *con* $\tau = \tau_0 = \tau_0\,(1 - \varkappa\rho_0)$ *es válida la estimación*

$$\| S \| \leqslant \bar{\rho}_0, \quad \bar{\rho}_0 = (1 - \bar{\xi})/(1 + \bar{\xi}), \tag{9'}$$

donde

$$\tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \varkappa = \frac{\gamma_3}{\sqrt{\gamma_1\gamma_2 + \gamma_3^2}}, \quad \bar{\xi} = \frac{1 - \varkappa}{1 + \varkappa} \cdot \frac{\gamma_2}{\gamma_1}.$$

Citemos la demostración del lema 4. Sea θ un número arbitrario del intervalo $(0, 1)$. Representemos el operador S en la forma siguiente:

$$S = E - \tau C = [\theta E - \tau C_0] + [(1 - \theta)\,E - \tau C_1],$$

donde $C_0 = 0{,}5\,(C + C^*)$ es la parte autoconjugada del operador C. Utilizando la desigualdad triangular, obtenemos la estimación para la norma del operador S:

$$\| S \| \leqslant \| \theta E - \tau C_0 \| + \| (1 - \theta)\,E - \tau C_1 \|. \tag{10}$$

Valoremos por separado la norma de cada operador. De (9) y de la igualdad $(C_0 x, x) = 0{,}5\,(Cx, x) + 0{,}5\,(C^* x, x) = (Cx, x)$ obtenemos, que γ_1 y γ_2 son las cotas del operador autoconjugado C_0:

$$\gamma_1 E \leqslant C_0 \leqslant \gamma_2 E, \quad \gamma_1 > 0.$$

Por analogía con el lema 2 obtenemos la siguiente estimación para la norma del operador $\theta E - \tau C_0$:

$$\| \theta E - \tau C_0 \| \leqslant \max (|\theta - \tau\gamma_1|, \ |\theta - \tau\gamma_2|) = \begin{cases} \theta - \tau\gamma_1, & 0 \leqslant \tau \leqslant \theta\tau_0, \\ \tau\gamma_2 - \theta, & \tau \geqslant \theta\tau_0. \end{cases}$$

Estimemos la norma del operador $(1 - \theta) E - \tau C_1$. Puesto que $(C_1 x, x) = 0$, entonces para todos los $x \in H$ obtenemos

$$\| ((1 - \theta) E - \tau C_1) x \|^2 = (1 - \theta)^2 \| x \|^2 + \tau^2 \| C_1 x \|^2 \leqslant$$
$$\leqslant ((1 - \theta)^2 + \tau^2 \| C_1 \|^2) \| x \|^2.$$

De esta expresión y de (9) se deduce la estimación $\| (1 - \theta) E - \tau C_1 \| \leqslant [(1 - \theta)^2 + \tau^2\gamma_3^2]^{1/2}$. Sustituyendo las estimaciones obtenidas en (10), tendremos

$$\| S \| \leqslant \begin{cases} \varphi_1 (0, \tau) = \theta - \tau\gamma_1 + \sqrt{(1-\theta)^2 + \tau^2\gamma_3^2}, & 0 \leqslant \tau \leqslant \tau_0\theta, \\ \varphi_2 (\theta, \tau) = \tau\gamma_2 - \theta + \sqrt{(1-\theta)^2 + \tau^2\gamma_3^2}, & \tau \geqslant \tau_0. \end{cases}$$

Elijamos ahora los parámetros τ y θ de la condición de mínimo de la estimación de la norma del operador S. Notemos que la función $\varphi_2 (0, \tau)$ crece monótonamente respecto de τ. Por eso para minimizar la norma del operador S es suficiente examinar el dominio $0 \leqslant \tau \leqslant \tau_0\theta$, $0 < \theta < 1$. En este dominio $\| S \| \leqslant \varphi_1 (\theta, \tau)$.

Investiguemos la función $\varphi_1 (\theta, \tau)$. Esta función crece monótonamente respecto de θ y por consiguiente su mínimo se alcanza para $\tau = \tau_0\theta$. Para este valor del parámetro τ tendremos

$$\| S \| \leqslant \varphi (\theta) = \varphi_1 (\theta, \tau_0\theta) = \theta (1 - \tau_0\gamma_1) +$$
$$+ \sqrt{(1-\theta)^2 + \tau_0^2\gamma_3^2\theta^2} = \theta\rho_0 + \sqrt{(1-\theta)^2 + \tau_0^2\gamma_3^2\theta^2}.$$

Así pues, es necesario mostrar que, $\min\limits_{0<\theta<1} \varphi (\theta) = \bar{\rho}_0$. Hallamos el mínimo de la función $\varphi (\theta)$. Hagamos la sustitución de variable, poniendo

$$\theta = (1 - x)/(1 + a^2), \quad x \in (-a^2, 1), \quad a^2 = \tau_0^2\gamma_3^2.$$

La función $\varphi (\theta)$ se escribe en la forma

$$\varphi (\theta) = \bar{\varphi} (x) = \frac{1}{\sqrt{1+a^2}} \left(\sqrt{x^2 + a^2} - \frac{\rho_0}{\sqrt{1+a^2}} x \right) + \frac{\rho_0}{1+a^2}. \qquad (11)$$

De aquí se ve, que es suficiente hallar el mínimo de la función

$$v (x) = \sqrt{x^2 + a^2} - \rho_0 x / \sqrt{1 + a^2}$$

en la región $-a^2 < x < 1$.. Calculando las derivadas de la función $v (x)$

$$v' (x) = \frac{x}{\sqrt{x^2+a^2}} - \frac{\rho_0}{\sqrt{1+a^2}}, \qquad v'' (x) = \frac{a^2}{(x^2+a^2)^{3/2}} > 0,$$

encontramos, que la ecuación $v' (x) = 0$ da el punto del mínimo de la función $v (x)$. Resolviendo la ecuación

$$\sqrt{\frac{x^2+a^2}{1+a^2}}=\frac{x}{\rho_0}, \tag{12}$$

hallamos el punto de mínimo buscado de la función $v(x)$:

$$x_0=a\rho_0/\sqrt{1+a^2-\rho_0^2}\in(0,\ 1),\quad \theta_0=(1-x_0)/(1+a^2).$$

Sustituyendo (12) en (11), hallemos el valor mínimo de la función $\varphi(\theta)$:

$$\varphi(\theta_0)=\sqrt{\frac{x_0^2+a^2}{1+a^2}}+\rho_0\frac{1-x_0}{1+a^2}=\frac{x_0}{\rho_0}+\theta_0\rho_0. \tag{13}$$

Queda por expresar x_0 y θ_0 mediante las magnitudes conocidas. Empleando las notaciones del lema 4, obtenemos

$$1-\rho_0^2=\tau_0^2\gamma_1\gamma_2,\quad x_0=\tau_0\gamma_2\rho_0/\sqrt{1-\rho_0^2+\tau_0^2\gamma_2^2}=\varkappa\rho_0. \tag{14}$$

De (12) encontramos

$$a^2=x_0^2(1-\rho_0^2)/(\rho_0^2-x_0^2),\quad 1+a^2=\rho_0^2(1-x_0^2)/(\rho_0^2-x_0^2).$$

Por eso

$$\theta_0\rho_0=\frac{1-x_0}{1+a^2}\rho_0=\frac{\rho_0^2-x_0^2}{\rho_0(1+x_0)}=\frac{\rho_0(1-\varkappa^2)}{1+\varkappa\rho_0}. \tag{15}$$

Sustituyamos (14) y (15) en (13):

$$\varphi(\theta)=\varkappa_0+\frac{\rho_0(1-\varkappa^2)}{1+\rho_0\varkappa}=\frac{\varkappa+\rho_0}{1+\rho_0\varkappa}=\frac{(1+\varkappa)-\xi(1-\varkappa)}{(1+\varkappa)+\xi(1-\varkappa)}=$$

$$=\frac{1-\bar{\xi}}{1+\bar{\xi}}=\bar{\rho}_0. \tag{16}$$

Hallemos ahora la expresión para el parámetro $\tau=\tau_0\theta_0$. Comparando (15) y (16), obtendremos

$$\theta_0\rho_0=\bar{\rho}_0-\varkappa. \tag{17}$$

Por otra parte, de (16) se puede expresar ρ_0 mediante $\bar{\rho}_0$ y $\varkappa$:

$$\rho_0=(\bar{\rho}_0-\varkappa)/(1-\varkappa\bar{\rho}_0).$$

Sustituyendo ρ_0 en (17), encontramos

$$\theta_0=1-\varkappa\bar{\rho}_0,\quad \tau=\tau_0(1-\varkappa\bar{\rho}_0).$$

El lema está demostrado.

Las desigualdades (9) pueden ser escritas en la forma siguiente:

$$\gamma_1(x,\ x)\leqslant(Cx,\ x)\leqslant\gamma_2(x,\ x),\quad (C_1x,\ C_1x)\leqslant\gamma_3^2(x,\ x),$$

$\gamma_1>0$.

Sustituyendo aquí $C = D^{-1/2}(DB^{-1}A)D^{-1/2}$ y $C_1 = 0{,}5\,D^{-1/2}(DB^{-1}A - (DB^{-1}A)^*)D^{-1/2}$, obtenemos las desigualdades

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0,$$

$$\left(D^{-1}\frac{DB^{-1}A-(DB^{-1}A)^*}{2}x,\ \frac{DB^{-1}A-(DB^{-1}A^*)}{2}x\right) \leqslant \gamma_3^2 (Dx, x). \tag{18}$$

Sustituyendo en (4) la estimación (9') para la norma del operador S, hallemos

$$\| z_n \|_D \leqslant \bar{\rho}_0^n \| z_0 \|_D. \tag{19}$$

TEOREMA 4. *Sean* γ_1, γ_2 y γ_3 *las constantes en las desigualdades* (18). *Para el valor* $\tau = \bar{\tau}_0 = \tau_0(1 - \varkappa\bar{\rho}_0)$ *del parámetro iterativo el método de iteración simple* (2) *converge en* H_D, *y para el error* z_n *tiene lugar la estimación* (19). *Para el número de iteraciones es cierta la estimación* $n \geqslant n_0(\varepsilon)$, *donde* $n_0(\varepsilon) = \ln \varepsilon / \ln \bar{\rho}_0$,

$$\tau_0 = \frac{2}{\gamma_1+\gamma_2}, \quad \bar{\rho}_0 = \frac{1-\bar{\xi}}{1+\bar{\xi}},$$

$$\bar{\xi} = \frac{1-\varkappa}{1+\varkappa}\cdot\frac{\gamma_1}{\gamma_2}, \quad \varkappa = \frac{\gamma_3}{\sqrt{\gamma_1\gamma_2+\gamma_3^2}}.$$

OBSERVACIÓN. Puesto que el número de iteraciones se determina por el valor $\bar{\xi}$, el cual se puede escribir en la forma

$$\bar{\xi} = (\sqrt{\gamma_1/\gamma_2 + (\gamma_3/\gamma_2)^2} - \gamma_2/\gamma^2)^2,$$

entonces el operador B se debe elegir de manera tal que el cociente $\xi = \gamma_1/\gamma_2$ sea máximo y γ_3/γ_2 mínimo.

Citamos ejemplos de elección del operador D. Si en calidad de D elegimos el operador A^*A o B^*B, entonces las desigualdades (18) se pueden escribir en la forma

$$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx) \leqslant \gamma_2 (Bx, Bx), \tag{20}$$
$$\| 0{,}5\,(AB^{-1} - (B^*)^{-1}A^*) \| \leqslant \gamma_3.$$

Realmente, para el caso $D = B^*B$ esta afirmación es evidente, y si $D = A^*A$, en (18) hay que hacer la sustitución $x = A^{-1}By$ y obtener las desigualdades (20).

Si el operador B es autoconjugado definido positivo y acotado en H, entonces en calidad de operador D se puede

tomar el operador B o $A^*B^{-1}A$. En este caso las desigualdades (18) son equivalentes a las desigualdades siguientes:

$$\gamma_1 B \leqslant A < \gamma_2 B, \quad \gamma_1 > 0, \tag{21}$$
$$(B^{-1}A_1 x,\ A_1 x) \leqslant \gamma_3^2 (Bx,\ x), \quad A_1 = 0{,}5\ (A - A^*).$$

En efecto, para $D = B$ las desigualdades (18) y (21) coinciden, y para $D = A^*B^{-1}A$ las desigualdades (21) se deducen de las desigualdades (18) después del cambio $x = A^{-1}By$ en (18).

3. Minimización de la norma del operador de resolución.

3.1 PRIMER CASO. En el punto 2 del § 4 fueron obtenidas las estimaciones para la norma del operador S^n, basados en la desigualdad $\| S^n \| \leqslant \| S \|^n$. Examinemos ahora otro método de obtención de estimaciones para $\| S^n \|$. Este método se basa en la estimación del radio numérico del operador.

Recordemos (véase el § 1 del cap. V), que se llama *radio numérico del operador* T, el cual actúa en el espacio de Hilbert H, la cantidad

$$\rho (T) = \sup_{\|z\|=1} | (Tz,\ z) |, \quad z \in H.$$

Para un operador lineal acotado T el radio numérico satisface las desigualdades

$$\mu (T) \| T \| \leqslant \rho (T) \leqslant \| T \|, \quad \rho (T^n) \leqslant [\rho (T)]^n \tag{22}$$

donde n es un número natural y $\mu (T) \geqslant 1/2$.

Utilizando el concepto de radio numérico de un operador, obtenemos dos estimaciones para la norma del operador S^n en dependencia del tipo de información a priori con respecto al operador C.

Examinemos el caso, cuando la información a priori está dada en forma de las constantes γ_1, γ_2 y γ_3:

$$\gamma_1 E \leqslant C < \gamma_2 E, \quad \| C_1 x \| \leqslant \gamma_3 \| x \|, \quad \gamma_1 > 0, \ x \in H. \tag{23}$$

Definiremos el espacio complejo $\widetilde{H}$ de la siguiente forma: consiste de los elementos del tipo $z = x + iy$, donde $x,\ y \in H$. El producto escalar en $\widetilde{H}$ se define por la fórmula

$$(z,\ w) = (x,\ u) + i\ (y,\ u) - i\ (x,\ v) + (y,\ v),$$
$$z = x + iy, \quad w = u + iv.$$

El operador lineal C, dado en H, lo definiremos sobre H de la siguiente forma:

$$Cz = Cx + iCy.$$

En virtud de la propiedad (22) para cualquier n entero es cierta la estimación

$$\| S^n \| \leqslant \frac{1}{\mu (S^n)} \rho (S^n) \leqslant 2\ [\rho (S)]^n,$$

por eso es suficiente estimar el radio numérico del operador S.

Tiene lugar

LEMA 5. *Sean dados* γ_1, γ_2 y γ_3 *en las desigualdades* (23). *Entonces para la norma del operador* $S = E - \tau C$ *en* H *con* $\tau = \min(\tau_0, \varkappa\tau_0)$ *es válida la estimación*

$$\| S^n \| \leqslant 2\rho^n,$$

donde

$$\rho^2 = \begin{cases} 1-\varkappa(1-\rho_0), & 0 < \varkappa \geqslant 1, \\ 1-(2-1/\varkappa)(1-\rho_0), & \varkappa \geqslant 1, \end{cases} \qquad \varkappa = \frac{\gamma_1(\gamma_1+\gamma_2)}{2(\gamma_1^2+\gamma_3^2)},$$

$$\tau_0 = 2/(\tau_1+\tau_2), \qquad \rho_0 = (1-\varepsilon)/(1+\xi), \qquad \xi = \gamma_1/\gamma_2.$$

Para la demostración del lema presentemos el operador C en forma de la suma $C = C_0 + C_1$, $C_0 = 0{,}5\,(C + C^*)$, $C_1 = 0{,}5\,(C - C^*)$. Calculemos el radio numérico del operador $S = E - \tau C$. Para cualquier $z \in \widetilde{H}$ obtenemos

$$(Sz, z) = (z, z) - \tau\,(C_0 z, z) - \tau\,(C_1 z, z).$$

En virtud de la autoconjugación del operador C_0 el producto escalar $(C_0 z, z)$ es número real. Como $C_1 = -C_1^*$, entonces $(C_1 z, z)$ es un número imaginario. Por lo tanto

$$|\,(Sz, z)\,|^2 = |(z, z) - \tau\,(C_0 z, z)|^2 + \tau^2\,|\,(C_1 z, z)\,|^2. \tag{24}$$

De las desigualdades (23) obtenemos

$$\gamma_1\,(z, z) \leqslant (C_0 z, z) \leqslant \gamma_2\,(z, z), \qquad \| C_1 z \| \leqslant \gamma_3\,|\,z\,|. \tag{25}$$

Sea $|\,z\,| = 1$. De (25) hallamos

$$|(z, z) - \tau\,(C_0 z, z)| \leqslant \max(|1-\tau\gamma_1|, |1-\tau\gamma_2|) =$$

$$= \begin{cases} 1-\tau\gamma_1, & 0 \leqslant \tau \leqslant \tau_0, \\ \tau\gamma_2 - 1, & \tau \geqslant \tau_0, \end{cases}$$

$$|\,(C_1 z, z)\,| \leqslant \| C_1 z \|\,\|\,z\,\| \leqslant \gamma_3.$$

Sustituyendo estas estimaciones en (24), obtendremos

$$\rho^2(S) = \sup_{\|z\|=1} |(Sz, z)|^2 \leqslant \begin{cases} \varphi_1(\tau) = (1-\tau\gamma_1)^2 + \tau^2\gamma_3^2, & 0 \leqslant \tau \leqslant \tau_0 \\ \varphi_2(\tau) = (1-\tau\gamma_2)^2 + \tau^2\gamma_3^2, & \tau \geqslant \tau_0. \end{cases}$$

Elijamos el parámetro τ de la condición de mínimo de la estimación del radio numérico operador S. Ya que la función $\varphi_2(\tau)$ crece según τ para $\tau \geqslant \tau_0$:

$$\varphi_2'(\tau) = 2\,]\tau\,(\gamma_2^2+\gamma_3^2) - \gamma_2] \geqslant 2\,\frac{\gamma_2(\gamma_2-\gamma_1)+2\gamma_3^2}{\gamma_1+\gamma_2} > 0,$$

entonces se debe determinar el mínimo de $\rho\,(S)$ en la región $\tau \leqslant \tau_0$ por τ, donde para $\rho\,(S)$ se cumple la estimación $\rho^2\,(S) \leqslant \varphi_1\,(\tau)$.

Investiguemos la función $\varphi_1\,(\tau)$. Puesto que

$$\varphi_1''\,(\tau) = 2\,(\gamma_1^2 + \gamma_3^2) > 0,$$

entonces igualando a cero la derivada

$$\varphi_1'\,(\tau) = 2\,[\tau\,(\gamma_1^2) + \gamma_3^2 - \gamma_1|,$$

hallamos el punto de extremo de la función $\varphi_1(\tau)$

$$\tau = \tau_1 = \frac{\gamma_1}{\gamma_1^2 + \gamma_3^2} = \tau_0 \varkappa.$$

Para $\tau \leqslant \tau_1$ la función $\varphi_1(\tau)$ decrece, y para $\tau \leqslant \tau_1$ crece. Por eso el mínimo de $\varphi_1(\tau)$ se alcanza en el punto $\tau = \tau_1$, si $\tau_1 \leqslant \tau_0$ y en el punto $\tau = \tau_0$, si $\tau_1 \geqslant \tau_0$. De esta forma hemos hallado el valor óptimo del parámetro τ: $\tau = \text{mín}(\tau_0, \tau_0 \varkappa)$. Con esto

$$\min_{\tau} \rho^2(S) \leqslant \begin{cases} \varphi_1(\tau_0), & \varkappa \geqslant 1, \\ \varphi_1(\tau_1), & 0 \leqslant \varkappa \leqslant 1. \end{cases}$$

Calculemos $\varphi_1(\tau_0)$ y $\varphi_1(\tau_1)$. De la definición de $\varkappa$ y de la igualdad $1 - \tau_0\gamma_1 = \rho_0$ obtenemos

$$\varkappa = \frac{\tau_0\gamma_1}{\tau_0^2\gamma_1^2 + \tau_0^2\gamma_3^2}, \quad \tau_0^2\gamma_3^2 = \frac{\tau_0\gamma_1}{\varkappa} - 1 - \tau_0^2\gamma_1^2 = \frac{1-\rho_0}{\varkappa} - (1-\rho_0)^2.$$

Así sucesivamente,

$$\varphi_1(\tau_0) = (1 - \tau_0\gamma_1)^2 + \tau_0^2\gamma_3^2 = \rho_0^2 + (1 - \rho_0)/\varkappa - (1 - \rho_0)^2 =$$
$$= 1 - (2 - 1/\varkappa)(1 - \rho_0),$$

$$\varphi_1(\tau_1) = (1 - \tau_1\gamma_1)^2 + \tau_1^2\gamma_3^2 = 1 - 2\tau_1\gamma_1 + \tau_1^2(\gamma_1^2 + \gamma_3^2) =$$
$$1 - \tau_1\gamma_1 = 1 - \varkappa\tau_0\gamma_1 = 1 - \varkappa(1 - \rho_0).$$

Así ha sido estimado el radio numérico. La estimación del lema se deduce de la desigualdad $\| S^n \| \leqslant 2[\rho(S)]^n$. El lema está demostrado.

Utilizando el lema 5, obtendremos una estimación para la norma del error z_n:

$$\| z_n \|_D \leqslant 2\rho^n \| z_0 \|_D. \tag{26}$$

TEOREMA 5. *Sean γ_1, γ_2 y γ_3 las constantes en las desigualdades (18). Para el valor $\tau = \min(\tau_0, \varkappa\tau_0)$ del parámetro iterativo, el método de iteración simple (2) converge en H_D, y para el error z_n tiene lugar la estimación (26). Para el número de iteraciones es cierta la estimación $n \geqslant n_0(\varepsilon)$, donde $n_0(\varepsilon) = \ln(0{,}5\varepsilon)/\ln\rho$, y $\varkappa$, ρ están definidos en el lema 5.*

Los ejemplos de elección del operador D y el tipo concreto de las desigualdades (18) se exponen en el punto 2.2.

3.2 SEGUNDO CASO. Aplicando el concepto de radio numérico de un operador, obtendremos otra estimación para la norma del operador S^n. Supongamos, que la información a priori está prefijada en forma de las constantes γ_1, γ_2 y γ_3 en las desigualdades

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad (C_1 x, C_1 x) \leqslant \gamma_3 (Cx, x), \quad \gamma_1 > 0. \tag{27}$$

Tiene lugar

LEMA 6. *Sean dadas γ_1, γ_2 y γ_3 en las desigualdades (27). Entonces para la norma del operador $S = E - \tau C$ en H para $\tau = \text{mín}(\tau_0^*, \varkappa\tau_0^*)$ es válida la estimación*

$$\| S^n \| \leqslant 2\rho^n,$$

donde

$$\rho^2 = \begin{cases} 1-(2\varkappa-1)\dfrac{1-\rho_0}{1+\rho_0}, & \dfrac{1}{2} \leqslant \varkappa \leqslant 1, \\ 1-\left(2-\dfrac{1}{\varkappa}\right)^2 \dfrac{1-\rho_0}{1+\rho_0}, & \varkappa \geqslant 1, \end{cases} \qquad \varkappa = \frac{\gamma_1+\gamma_2+\gamma_3}{2(\gamma_1+\gamma_3)},$$

$$\tau_0^* = 2/(\gamma_1+\gamma_2+\gamma_3), \quad \rho_0 = (1-\xi)/(1+\xi), \quad \xi = \gamma_1/\gamma_2.$$

En efecto, representando el operador C en la forma $C = C_0 + C_1$, donde $C_0 = 0{,}5\,(C + C^*)$ y $C_1 = 0{,}5\,(C - C^*)$, obtenemos

$$|\,(Sz,\ z)\,|^2 = |(z,\ z) - \tau\,(C_0 z,\ z)|^2 + \tau^2\,|\,(C_1 z,\ z)\,|^2.$$

De la desigualdad de Cauchy-Buniakovski y de las condiciones del lema hallemos

$$|\,(C_1 z,\ z)\,|^2 \leqslant (C_1 z,\ C_1 z)\,(z,\ z) \leqslant \gamma_3\,(C_0 z,\ z)\,(z,\ z).$$

Como para cualquier $z \in \widetilde{H}$ tienen lugar las desigualdades

$$\gamma_1\,(z,\ z) \leqslant (C_0 z,\ z) \leqslant \gamma_2\,(z,\ z), \qquad \gamma_1 > 0,$$

entonces de las tres relaciones anteriores obtenemos la siguiente estimación para el radio numérico del operador S:

$$\rho^2\,(S) \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \varphi\,(t), \text{ donde } \varphi\,(t) = (1-\tau t)^2 + \tau^2 \gamma_3 t.$$

Investiguemos la función $\varphi\,(t)$. Esta función puede tomar el valor máximo solamente en los extremos del segmento $[\gamma_1,\ \gamma_2]$. Por lo tanto

$$\rho^2\,(S) \leqslant \begin{cases} \varphi_1\,(\tau) = (1-\tau\gamma_1)^2 + \tau^2\gamma_1\gamma_3, & 0 \leqslant \tau \leqslant \tau_0^*, \\ \varphi_2\,(\tau) = (1+\tau\gamma_2)^2 + \tau^2\gamma_2\gamma_3, & \tau \geqslant \tau_0^*. \end{cases}$$

Elijamos el parámetro τ de la condición de mínimo de la estimación para $\rho\,(S)$. Ya que la función $\varphi_2\,(\tau)$ crece según τ para $\tau \geqslant \tau_0^*$:

$$\varphi_2'\,(\tau) = 2\gamma_2\,[\tau\,(\gamma_2+\gamma_3) - 1] \geqslant 2\gamma_2\,\frac{\gamma_2-\gamma_1+\gamma_3}{\gamma_2+\gamma_1+\gamma_3} > 0,$$

entonces se debe buscar el mínimo de $\rho\,(S)$ en la región $\tau \leqslant \tau_0^*$, donde para $\rho\,(S)$ se cumple la estimación $\rho^2\,(S) \leqslant \varphi_1\,(\tau)$.

La función $\varphi_1\,(\tau)$ para $\tau = \tau_1 = 1/(\gamma_1 + \gamma_3) = \varkappa\tau_0^*$ alcanza el valor mínimo, y al mismo tiempo decrece para $\tau \leqslant \tau_1$ y crece para $\tau \geqslant \tau_1$. Por lo tanto el valor mínimo de $\varphi_1\,(\tau)$ sobre el segmento $[0,\ \tau_0^*]$ se alcanza en el punto $\tau = \tau_1$, si $\tau_1 \leqslant \tau_0^*$ y en el punto $\tau = \tau_0^*$ si $\tau_1 \geqslant \tau_0^*$.

De esta forma, está hallado el valor óptimo del parámetro τ:

$$\tau = \text{mín}\,(\tau_0^*,\ \varkappa\tau_0^*).$$

Con esto

$$\min_{\tau} \rho^2\,(S) \leqslant \begin{cases} \varphi_1\,(\tau_0^*), & \varkappa \geqslant 1, \\ \varphi_1\,(\tau_1), & \varkappa \leqslant 1. \end{cases}$$

Calculemos $\varphi_1(\tau_0^*)$ y $\varphi_1(\tau_1)$. Unos cálculos no complicados dan $\tau_0^* = (2 - 1/\varkappa)/\gamma_2$, $\gamma_3 = 1/(\varkappa\tau_0^*) - \gamma_1$. Empleando estas relaciones, obtenemos

$$\varphi_1(\tau_0^*) = (1 - \tau_0^*\gamma_1)^2 + (\tau_0^*)^2\gamma_1\gamma_3 = 1 - 2\tau_0^*\gamma_1 + \tau_0^*\gamma_1/\varkappa =$$
$$= 1 - (2 - 1/\varkappa)^2\gamma_1/\gamma_2 = (1 - (2 - 1/\varkappa)^2(1 - \rho_0)/(1 + \rho_0).$$

A continuación

$$\varphi_1(\tau_1) = (1 - \tau_0^*\varkappa\gamma_1)^2 + (\tau_0^*)^2\varkappa^2\gamma_1\gamma_3 = 1 - \tau_0^*\varkappa\gamma_1 =$$
$$= 1 - (2\varkappa - 1)\gamma_1/\gamma_2 = 1 - (2\varkappa - 1)(1 - \rho_0)/(1 + \rho_0).$$

Así pues está estimado el radio numérico. La estimación del lema se deduce de la desigualdad $\| S^n \| \leqslant 2[\rho(S)]^n$. El lema está demostrado.

Sustituyendo $C = D^{-1/2}(DB^{-1}A)D^{-1/2}$ y $C_1 = 0{,}5D^{-1/2} \times (DB^{-1}A - (DB^{-1}A)^*)D^{-1/2}$ en las desigualdades (27), obtendremos las desigualdades

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0, \tag{28}$$
$$\left(D^{-1}\frac{DB^{-1}A - (DB^{-1}A)^*}{2}x,\ \frac{DB^{-1}A - (DB^{-1}A)^*}{2}x\right) \leqslant \gamma_3(DB^{-1}Ax,\ x).$$

TEOREMA 6. *Sean γ_1, γ_2 y γ_3 las constantes en las desigualdades (28). Para el valor $\tau = \min(\tau_0^*, \varkappa\tau_0^*)$ del parámetro iterativo el método de iteración simple (2) converge en H_D, y para el error z_n tiene lugar la estimación (26). Para el número de iteraciones es cierta la estimación $n \geqslant n_0(\varepsilon)$, donde*

$$n_0(\varepsilon) = \ln 0{,}5\varepsilon/\ln\rho,$$

y $\varkappa$, ρ y τ_0^* están definidos en el lema 6.

Mostremos el tipo de las desigualdades (28) por algunos ejemplos de elección del operador D. Si en calidad del operador D tomamos el operador A^*A ó B^*B, entonces las desigualdades (28) se pueden escribir en la forma siguiente:

$$\gamma_1(Bx,\ Bx) \leqslant (Ax,\ Bx) \leqslant \gamma_2(Bx,\ Bx), \quad \gamma_1 > 0,$$
$$\| 0{,}5(AB^{-1} - (B^*)^{-1}A^*)x \|^2 \leqslant \gamma_3(A^*x,\ B^{-1}x). \tag{29}$$

En efecto, las desigualdades (29) se desprenden directamente de (28) después de la sustitución de $D = B^*B$ en (29). Para el caso $D = A^*A$ en (28) es suficiente hacer la sustitución $x = A^{-1}By$.

Si el operador B es autoconjugado, definido positivo y acotado en H, entonces en calidad de operador D se pueden tomar los operadores B o $A^*B^{-1}A$. En este caso las desigualdades (28) tendrán la forma

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0,$$
$$(B^{-1}A_1x,\ A_1x) \leqslant \gamma_3(Ax,\ x), \quad A_1 = 0{,}5(A - A^*). \tag{30}$$

Señalemos, que para el caso de $D = A^*B^{-1}A$ las desigualdades (30) se deducen de (28) después del cambio indicado más arriba.

4. **Método de simetrización de una ecuación.** Para resolver la ecuación $Au = f$ con el operador no autoconjugado A se utiliza el bien conocido método de *simetrización de una ecuación*. En lugar de la ecuación inicial se examina la ecuación simetrizada

$$\tilde{A}u = \tilde{f}, \quad \tilde{A} = A^*A, \quad \tilde{f} = A^*f, \tag{31}$$

la cual se obtiene de la ecuación inicial multiplicando de la izquierda por el operador conjugado a A. En el álgebra esta transformación de la ecuación se llama *primera transformación de Gauss*.

Para la resolución aproximada de la ecuación (31) examinaremos el esquema implícito de dos capas

$$\tilde{B}\frac{y_{k+1}-y_k}{\tau_{k+1}} + \tilde{A}y_k = \tilde{f}, \quad k = 0, 1, \ldots, \quad y_0 \in H, \tag{32}$$

con el operador autoconjugado y definido positivo $\tilde{B}$. En calidad de operador D elegiremos los operadores $\tilde{B}$ o $\tilde{A} = A^*A$. En este caso el operador $D\tilde{B}^{-1}\tilde{A}$ es autoconjugado en H y por eso los parámetros iterativos τ_k pueden ser elegidos por las fórmulas del método de Chebishev, investigado en el § 2. La información a priori pasa este método, en el caso de los operadores D indicados, tiene la forma de las constantes de equivalencia energética de los operadores $\tilde{B}$ y $\tilde{A} = A^*A$

$$\gamma_1\tilde{B} \leqslant A^*A \leqslant \gamma_2\tilde{B}, \quad \gamma_1 > 0.$$

La estimación de la velocidad de convergencia del método de Chebishev (32) y las fórmulas para los parámetros iterativos están expuestos en el teorema 1.

§ 5. Ejemplos de aplicación de los métodos iterativos

1. **Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo.** Para ilustrar la aplicación de los métodos iterativos de dos capas construidos en este capítulo examinaremos la resolución del problema de Dirichlet de diferencias para las ecuaciones lineales elípticas

de segundo orden. Interpretaremos un problema de diferencias como una ecuación operacional

$$Au = f \tag{1}$$

en un espacio de dimensión finita de funciones reticulares. Serán examinados los métodos de Chebishev explícito e implícito, y también el método de iteración simple.

Comenzaremos el exámen de los ejemplos por *el problema de Dirichlet para la ecuación de Poisson en un rectángulo.* Supongamos que en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$ con frontera Γ se exige hallar la solución de la ecuación de Poisson

$$Lu = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -f(x),\ x = (x_1, x_2) \in G, \tag{2}$$

que toma los valores prefijados sobre la frontera Γ

$$u(x) = g(x),\ x \in \Gamma \tag{3}$$

El problema de Dirichlet de diferencias correspondiente a (2), (3) sobre la red rectangular

$$\bar{\omega} = \{x_{ij} = (ih_1, jh_2) \in \bar{G},\quad 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\quad h_\alpha = l_\alpha/N,\quad \alpha = 1, 2\}$$

tiene la forma

$$\Lambda y = \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha x_\alpha} = -\varphi(x),\ x \in \omega,\ y(x) = g(x),\ x \in \gamma \tag{4}$$

donde $\gamma = \{x_{ij} \in \Gamma\}$ es la frontera de la red $\bar{\omega}$, y

$$y_{\bar{x}_1 x_1} = \frac{1}{h_1^2}(y(i+1, j) - 2y(i, j) + y(i-1, j)),$$

$$y_{\bar{x}_2 x_2} = \frac{1}{h_2^2}(y(i, j+1) - 2y(i, j) + y(i, j-1)),$$

$$y(i, j) = y(x_{ij}).$$

En el § 2 del cap. V fue mostrado, que el problema de diferencias (4) puede ser reducido a una ecuación operacional (1), para la cual el operador A se define de la siguiente forma: $Ay = -\Lambda_y^0$, donde $y \in H$, $\overset{\circ}{y} \in \overset{\circ}{H}$ y $\overset{\circ}{y}(x) = y(x)$ para $x \in \omega$. Aquí $\overset{\circ}{H}$ es el conjunto de las funciones reticulares definidas sobre $\bar{\omega}$ y que se anulan sobre γ, y H es el espacio de las funciones reticulares definidas sobre ω con el pro-

ducto escalar $(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2$. El segundo miembro f de la ecuación (1) se diferencia del miembro derecho φ de la ecuación en diferencias (4) solamente en los nodos fronterizos:

$$f(x) = \varphi(x) + \varphi_1(x)/h_1^2 + \varphi_2(x)/h_2^2,$$

$$\varphi_1(u) = \begin{cases} g(0, x_2), & x_1 = h_1, \\ 0, & 2h_1 \leqslant x_1 \leqslant l_1 - 2h_1 \\ g(l_1, x_2), & x_1 = l_1 - h_1, \end{cases}$$

$$\varphi_2(x) = \begin{cases} g(x_1, 0), & x_2 = h_2, \\ 0, & 2h_2 \leqslant x_2 \leqslant l_2 - 2h_2, \\ g(x_1, l_2), & x_2 = l_2 - h_2. \end{cases}$$

De esta forma, el problema de contorno de diferencias (4) está reducido a la ecuación operacional (1) en el espacio de Hilbert de dimensión finita H.

Para la resolución aproximada de la ecuación (1) examinemos EL MÉTODO DE CHEBISHEV EXPLÍCITO ($B = E$):

$$B \frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = f, \quad k = 0, 1, \ldots, \quad y_0 \in H \tag{5}$$

$$\tau_k = \frac{\tau_0}{1 + \rho_0 \rho_k}, \quad \mu_k \in \mathfrak{M}_n^* = \left\{ -\cos \frac{(2i-1)\pi}{2n}, \right.$$

$$\left. i = 1, 2, \ldots, n \right\}, \tag{6}$$

$$k = 1, 2, \ldots, n,$$

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \ln(0{,}5\varepsilon)/\ln \rho_1. \tag{7}$$

En el § 2 del cap. V se mostró, que el operador A aquí definido es autoconjugado en H y sus cotas γ_1 y γ_2 coinciden con los valores propios mínimo y máximo del operador de diferencias Λ, es decir

$$\gamma_1 E \leqslant A \leqslant \gamma_2 E, \quad \gamma_1 > 0, \tag{8}$$

donde

$$\gamma_1 = \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2 l_\alpha}, \quad \gamma_2 = \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2 l_\alpha}. \tag{9}$$

Los operadores A y $B = E$ son autoconjugados y definidos positivos en H. Por eso de los ejemplos examinados en el punto 3 del § 2 se deduce, que γ_1 y γ_2 de (8) son las

constantes para el método de Chebishev (5)-(7), si en calidad de D se ha elegido uno de los operadores E, A ó A^2. Entonces en las fórmulas (6), (7)

$$\tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

donde γ_1 y γ_2 están definidas en (9).

Ya que $\gamma_1 = O(1)$ y $\gamma_2 = O(1/h_1^2 + 1/h_2^2)$, entonces $\xi = O(|h|^2)$, donde $|h|^2 = h_1^2 + h_2^2$. Por lo tanto, para el ejemplo examinado la estimación asintótica del número de iteraciones $n_0(\varepsilon)$ tiene la forma

$$n_0(\varepsilon) = O\left(\frac{2}{|h|} \ln \frac{2}{\varepsilon}\right).$$

En el caso particular, cuando $\overline{G}$ es un cuadrado con lado l ($l_1 = l_2 = l$) y la red $\overline{\omega}$ es cuadrada ($h_1 = h_2 = h = 1/N$), tenemos

$$\gamma_1 = \frac{8}{h^2} \operatorname{sen}^2 \frac{\pi h}{2l}, \quad \gamma_2 = \frac{8}{h^2} \cos^2 \frac{\pi h}{2l}, \quad \xi = \operatorname{tg}^2 \frac{\pi h}{2l},$$

$$\tau_0 = \frac{h^2}{4}, \quad \rho_0 = \cos \frac{\pi h}{l}, \quad \rho_1 = \frac{1 - \operatorname{sen} \frac{\pi h}{l}}{\cos \frac{\pi h}{l}},$$

$$n_0(\varepsilon) \approx \frac{1}{\pi h} \ln \frac{2}{\varepsilon} \approx 0{,}32\, N \ln \frac{2}{\varepsilon}. \qquad (10)$$

De esta forma, el número n de iteraciones es proporcional al número de nodos N por una dirección. Señalemos, que el número de incógnitas en el problema (4) es igual a $M = (N-1)^2$, es decir el número de iteraciones es proporcional a la raíz cuadrada del número de incógnitas.

Utilizando la definición del operador A y el segundo miembro f, se puede escribir el esquema iterativo operacional (5) para $B = E$, en forma del siguiente esquema de diferencias:

$$y_{k+1} = y_k + \tau_{k+1}(\Lambda y_k + \varphi), \quad x \in \omega, \quad y_{k|\gamma} = g,$$
$$k = 0, 1, \ldots .$$

Sustituyendo aquí (4), obtendremos las fórmulas de cálculo

$$y_{k+1}(i, j) = \left(1 - \frac{\tau_{k+1}}{\tau_0}\right) y_k(i, j) + \\ + \tau_{k+1}\left[\frac{y_k(i+1, j) + y_k(i-1, j)}{h_1^2} + \right. \\ \left. + \frac{y_k(i, j+1) + y_k(i, j-1)}{h_2^2} + \varphi(i, j)\right],$$

$$1 \leqslant i \leqslant N_1 - 1,\ 1 \leqslant j \leqslant N_2 - 1.$$

La aproximación inicial y_0 es una función reticular arbitraria sobre ω, la cual toma sobre la frontera γ los valores prefijados $y_0(x) = g(x)$ para $x \in \gamma$.

Calculemos el número de operaciones $Q(\varepsilon)$, que conviene efectuar para obtener una solución aproximada del problema de diferencias (4) con exactitud ε mediante el método de Chebishev (5)-(7).

Considerando prefijados los parámetros iterativos τ_k, obtenemos, que para calcular y_{k+1} en un nodo de la red ω se necesitan nueve operaciones aritméticas. Ya que el número de nodos interiores de la red ω es igual a $M = (N_1 - 1) \times (N_2 - 1)$, entonces en la realización de un paso iterativo se exige $Q_0 \approx 9N_1N_2$ operaciones. Por lo tanto $Q(\varepsilon) = nQ_0 \approx 9nN_1N_2$, donde n es el número de iteraciones.

Para el caso particular examinado más arriba el número de iteraciones n está determinado en (10), y por consiguiente, para este ejemplo obtenemos

$$Q(\varepsilon) \approx 2{,}9\, N^3 \ln(2/\varepsilon).$$

Para resolver la ecuación (1) examinemos ahora EL METODO DE ITERACIÓN SIMPLE. El esquema iterativo del método de iteración simple tiene la forma (5), y los parámetros iterativos τ_k y el número de iteraciones n se determinan por las fórmulas del teorema 2:

$$\tau_k \equiv \tau_0 = \frac{2}{\gamma_1 + \gamma_2},\quad n \geqslant n_0(\varepsilon) = \frac{\ln \varepsilon}{\ln \rho_0},$$

$$\rho_0 = \frac{1-\xi}{1+\xi},\quad \xi = \frac{\gamma_1}{\gamma_2}, \qquad (11)$$

donde γ_1 y γ_2 están dadas en (9). De (9) y (11) obtenemos una estimación asintótica respecto de h del número de iteraciones para el método de iteración simple $n_0(\varepsilon) =$

$= O\left(\frac{1}{|h|^2} \ln \frac{1}{\varepsilon}\right)$. Para el caso particular examinado más arriba halla os

$$n_0(\varepsilon) \approx \frac{2l^2}{\pi^2 h^2} \ln \frac{1}{\varepsilon} \approx 0{,}2N^2 \ln \frac{1}{\varepsilon}, \qquad (12)$$

es decir el número de iteraciones para el método de iteración simple es proporcional al cuadrado del número de nodos N en una dirección (o proporcional al número de incógnitas en la ecuación).

Comparando las estimaciones para el número de iteraciones del método de Chebishev (10) y del método de iteración simple (12), obtenemos, que el método de iteración simple exige un número considerablemente mayor de iteraciones que el método de Chebishev. Para comparar estos métodos sobre reales mostremos el valor exacto del número de iteraciones para el caso particular indicado en dependencia del número N de nodos por una dirección para $\varepsilon = 10^{-4}$ (primero se indica el número de iteraciones para el método de Chebishev):

$N = 32 \quad n = 101 \quad n = 1909$

$N = 64 \quad n = 202 \quad n = 7642$

$N = 128 \quad n = 404 \quad n = 30577.$

Expongamos las fórmulas de cálculo del método de iteración simple para el caso particular examinado:

$$y_{k+1}(i, j) = \frac{1}{4}[y_k(i+1, j) + y_k(i-1, j) + y_k(i, j+1) + y_k(i, j-1)] + \frac{h^2}{4}\varphi(i, j).$$

La dependencia muy fuerte del número de iteraciones del método de iteración simple del número de nodos N de la red es la causa de que en la actualidad este método casi no se utiliza para resolver ecuaciones elípticas reticulares.

2. Problema de Dirichlet de diferencias para la ecuación de Poisson en una región arbitraria. Supongamos que en una región acotada arbitraria G con frontera Γ se exige hallar la solución de la ecuación de Poisson

$$Lu = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -f(x), \quad x = (x_1, x_2) \in G, \qquad (13)$$

que toma sobre la frontera Γ los valores dados

$$u(x) = g(x), \quad x \in \Gamma \qquad (14)$$

Para simplificar examinemos el caso, cuando la intersección de la región G con una recta, que pasa por el punto $x \in G$, siendo paralela al eje de coordenadas, consiste de un intervalo.

Cubramos el plano con un retículo formado por la intersección de rectas paralelas a los ejes coordenados y situadas a la distancia h la una de la otra.

A los puntos x_{ij} del retículo, pertenecientes a G, los llamaremos *nodos* de la red $\omega = \{x_{ij} \in G\}$. Mediante $\Delta_\alpha(x_{ij})$ designaremos el intervalo formado al intersecar G con la recta trazada por el punto $x_{ij} \in \omega$ paralelamente al eje coordenado Ox_α, $\alpha = 1, 2$. Llamaremos *nodos frontera* en la dirección x_α los extremos de este intervalo. El conjunto de todos los nodos frontera en la dirección x_α lo designaremos mediante γ_α, y mediante $\gamma = \gamma_1 \cup \gamma_2$ designaremos la frontera de la región reticular. El conjunto de los nodos interiores y de frontera forma una red $\bar{\omega} = \omega \cup \gamma$ en la región $\bar{G}$.

Examinemos uno de los intervalos Δ_α. El conjunto de los nodos $x_{ij} \in \omega$, que están en este intervalo, lo designaremos mediante $\omega_\alpha(x_\beta)$, $\beta = 3 - \alpha$, $\alpha = 1, 2$. Mediante $\omega_\alpha^+(x_\beta)$ designaremos el conjunto que consiste de los nodos de $\omega_\alpha(x_\beta)$ y del extremo derecho del intervalo Δ_α. Definamos $\omega_\alpha(x_\beta)$ como el conjunto, que consiste de los nodos $\omega_\alpha(x_\beta)$ y de los extremos del intervalo Δ_α. Designemos mediante $x_{ij}^{(+1\alpha)}$ y $x_{ij}^{(-1\alpha)}$ los nodos más cercanos al punto $x_{ij} \in \omega_\alpha(x_\beta)$ respectivamente a la derecha y a la izquierda y pertenecientes a $\omega_\alpha(x_\beta)$.

Llamaremos *pasos* $h_\alpha^\pm(x_{ij})$ de la red ω en el punto $x_{ij} \in \omega$ la distancia entre los nodos x_{ij} y $x_{ij}^{(\pm 1\alpha)} \in \omega$. Señalemos, que si todos los cuatro nodos $x_{ij}^{(\pm 1\alpha)}$ vecinos a x_{ij} pertenecen a ω entonces los pasos $h_\alpha^\pm$ son iguales al paso fundamental h del retículo. En los nodos fronterizos $h_\alpha^\pm \leqslant h$. Entre los pasos h_μ^+ y h_α^- tiene lugar la relación

$$h_\alpha^+(x_{ij}) = h_\alpha^-(x)_{ij}^{(+1\alpha)}.$$

Al problema (13), (14) le pondremos en correspondencia sobre la red $\bar{\omega}$ el problema de contorno de diferencias

$$\Lambda y = \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha \hat{x}_\alpha} = -\varphi(x), \quad x \in \omega, \quad y(x) = g(x), \quad x \in \gamma, \qquad (15)$$

donde

$$y_{\bar{x}_\alpha} = \frac{1}{h_\alpha^-}(y - y^{-1_\alpha}),\ y_{x_\alpha} = \frac{1}{h_\alpha^+}(y^{+1_\alpha} - y),$$

$$y_{\hat{x}_\alpha} = \frac{1}{h}(y^{+1_\alpha} - y),\ y_{\bar{x}_\alpha \hat{x}_\alpha} = \frac{1}{h}\left(\frac{y^{+1_\alpha} - y}{h_\alpha^+} - \right.$$

$$\left. - \frac{y - y^{-1_\alpha}}{h_\alpha^-}\right),\ y^{\pm 1_\alpha} = y(x^{(\pm 1_\alpha)}),\ \alpha = 1,\ 2.$$

El problema de contorno de diferencias (15) se reduce a la ecuación operacional (1) y el operador A se define de la misma forma que en el punto 1. El producto escalar en H se da de la siguiente forma:

$$(u,\ v) = \sum_{x \in \omega} u(x)\, v(x)\, h^2.$$

Introduzcamos algunas notaciones, que utilizaremos en lo sucesivo. Definamos para las funciones reticulares, prefijadas sobre $\overline{\omega}$, los productos escalares según las fórmulas:

$$(u,\ v)_{\omega_\alpha} = \sum_{x_\alpha \in \omega_\alpha (x_\beta)} u(x)\, v(x)\, h,$$

$$(u,\ v)_{\omega_\alpha^+} = \sum_{x_\alpha \in \omega_\alpha^+ (x_\beta)} u(x)\, v(x)\, h_\alpha^-(x),$$

$$(u,\ v)_\alpha = ((u,\ v)_{\omega_\alpha^+},\ 1)_{\omega_\beta},\ \beta = 3 - \alpha,\ \alpha = 1,\ 2.$$

Utilizando estas notaciones, el producto escalar en H se puede escribir en la forma

$$(u,\ v) = ((u,\ v)_{\omega_1},\ 1)_{\omega_2} = ((u,\ v)_{\omega_2},\ 1)_{\omega_1}. \qquad (16)$$

De la definición del operador A, obtenemos que

$$(Au,\ v) = -(\Lambda \overset{\circ}{u},\ \overset{\circ}{v}) = -((\overset{\circ}{u}_{\bar{x}_1 \hat{x}_1},\ \overset{\circ}{v})_{\omega_1},\ 1)_{\omega_2} -$$

$$- ((\overset{\bullet}{u}_{\bar{x}_2 \hat{x}_2},\ \overset{\bullet}{v})_{\omega_2},\ 1)_{\omega_1},$$

y como en virtud de las fórmulas de Green de diferencias tiene lugar la igualdad (véase el punto 3 del § 2 cap. V)

$$-(\overset{\circ}{u}_{\bar{x}_\alpha \hat{x}_\alpha},\ \overset{\circ}{v})_{\omega_\alpha} = (\overset{\circ}{u}_{\bar{x}_\alpha},\ \overset{\bullet}{v}_{\bar{x}_\alpha})_{\omega_\alpha^+} = -(\overset{\circ}{u},\ \overset{\circ}{v}_{\bar{x}_\alpha \hat{x}_\alpha})_{\omega_\alpha},$$

entonces de aquí se infieren las igualdades

$$(Au,\ v) = (u,\ Av),$$

$$(Au,\ u) = \sum_{\alpha=1}^{2} (\overset{\circ}{u}_{\overline{x}_\alpha},\ 1)_\alpha,\ u \in H,\ \overset{\circ}{u} \in \overset{\circ}{H}. \tag{17}$$

La primera de estas igualdades demuestra la autoconjugación del operador A en H.

Para la resolución aproximada de la ecuación (1) examinemos EL MÉTODO DE CHEBISHEV IMPLÍCITO

$$B\,\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f,\quad k = 0,\ 1,\ \ldots,\quad y_0 \in H,$$

donde en calidad de operador B tomamos el operador diagonal fácilmente inversible

$$By = (b_1 + b_2)\,y,\ b_\alpha(x) = \frac{1}{h}\left(\frac{1}{h_\alpha^+(x)} + \frac{1}{h_\alpha^-(x)}\right),\ \alpha = 1,\ 2. \tag{18}$$

Aclaremos la elección del operador B. Si se interpreta la ecuación (1) como un sistema de ecuaciones lineales algebraicas de la matriz $\mathscr{A}$ correspondiente al operador A, entonces la matriz $\mathscr{B}$, correspondiente al operador B, es la parte diagonal de la matriz $\mathscr{A}$.

Puesto que los operadores A y B son autoconjugados y definidos positivos en H, entonces γ_1 y γ_2, que entran en las condiciones (6) y (7), son las constantes de equivalencia energética de los operadores A y B:

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B,\quad \gamma_1 > 0,$$

si en calidad de D es elegido uno de los operadores A, B ó $AB^{-1}A$.

Hallemos las estimaciones de γ_1 y γ_2. Primeramente mostremos que tiene lugar la igualdad

$$\gamma_1 + \gamma_2 = 2. \tag{19}$$

En efecto, sea $u(x)$ una función reticular arbitraria en H. Examinemos la función $v(x)$, la cual definiremos de la siguiente forma:

$$v(x_{ij}) = (-1)^{i+j}\,u(x_{ij}),\quad x_{ij} \in \omega.$$

Calculemos el valor del operador en diferencias $\Lambda \overset{\circ}{v}$ en el punto x_{ij}. Obtenemos

$$\Lambda \overset{\circ}{v}(i, j) = \sum_{\alpha=1}^{2} \frac{1}{h} \left(\frac{\overset{\circ}{v}{}^{+1_\alpha} - \overset{\circ}{v}}{h_\alpha^+} + \frac{\overset{\circ}{v} - \overset{\circ}{v}{}^{-1_\alpha}}{h_\alpha^-} \right)_{x=x_{ij}} =$$

$$= -(-1)^{i+j} \sum_{\alpha=1}^{2} \frac{1}{h} \left(\frac{\overset{\circ}{u}{}^{+1_\alpha} - \overset{\circ}{u}}{h_\alpha^+} - \frac{\overset{\circ}{u} - \overset{\circ}{u}{}^{-1_\alpha}}{h_\alpha^-} \right)_{x=x_{ij}} -$$

$$- 2(-1)^{i+j} (b_1(x_{ij}) + b_2(x_{ij})) \overset{\circ}{u}(x_{ij}).$$

Por lo tanto,

$$Av(i, j) = -\Lambda \overset{\circ}{v}(i, j) = (-1)^{i+j} (2B - A) u(i, j).$$

Después, ya que

$$\gamma_1 = \min_{u \neq 0} \frac{(Au, u)}{(Bu, u)}, \quad (Av, v) = 2(Bu, u) - (Au, u),$$

$$(Bv, v) = (Bu, u),$$

entonces

$$\gamma_2 = \max_{v \neq 0} \frac{(Av, v)}{(Bv, v)} = 2 - \min_{u \neq 0} \frac{(Au, u)}{(Bu, u)} = 2 - \gamma_1,$$

La afirmación está demostrada.

Utilizando la relación (19), obtendremos, que en las fórmulas (6)

$$\tau_0 = 2/(\gamma_1 + \gamma_2) = 1, \quad \rho_0 = (\gamma_2 - \gamma_1)/(\gamma_2 + \gamma_1) = 1 - \gamma_1.$$

Por consiguiente, para calcular los parámetros iterativos τ_k es suficiente hallar la estimación para γ_1. Del lema 13 del § 2 cap. V obtenemos, que para cualquier función reticular $\overset{\circ}{y} \in \overset{\circ}{H}$ tiene lugar la desigualdad

$$(b_\alpha \overset{\circ}{y}, \overset{\circ}{y})_{\omega_\alpha} \leqslant \varkappa_\alpha (\overset{\circ}{y}{}^2_{\overline{x}_\alpha}, 1)_{\omega_\alpha^+}, \quad \alpha = 1, 2, \tag{20}$$

donde $\varkappa_\alpha = \varkappa_\alpha(x_\beta) = \max\limits_{x_\alpha \in \omega_\alpha(x_\beta)} y^\alpha(x)$, y $v^\alpha(x)$

es la solución del siguiente problema de contorno tripuntual:

$$\begin{aligned} &v^\alpha_{\overline{x}_\alpha \hat{x}_\alpha} = -b_\alpha(x), \quad x_\alpha \in \omega_\alpha(x_\beta), \\ &v^\alpha(x) = 0, \quad x_\alpha \in \gamma_\alpha \end{aligned} \tag{21}$$

Dividiendo la desigualdad (20) por $\varkappa_\alpha$ y sumando según ω_β obtendremos

$$\left(\frac{b_\alpha}{\varkappa_\alpha}\overset{\circ}{y},\ \overset{\circ}{y}\right)\leqslant((\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1)_{\omega_\alpha^+},\ 1)_{\omega_\beta}=(\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1)_\alpha,\ \alpha=1,\ 2.$$

Sumando estas desigualdades, hallaremos

$$\left(\sum_{\alpha=1}^{2}\frac{b_\alpha}{\varkappa_\alpha}\overset{\circ}{y},\ \overset{\circ}{y}\right)\leqslant\sum_{\alpha=1}^{2}(\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1)_\alpha. \tag{22}$$

De (17), (18) y (22) se deduce, que en calidad de γ_1 se puede tomar la magnitud

$$\gamma_1=\min_{x\in\omega}\frac{1}{b_1(x)+b_2(x)}\sum_{\alpha=1}^{2}\frac{b_\alpha(x)}{\varkappa_\alpha(x_\beta)}. \tag{23}$$

Queda por calcular $\varkappa_\alpha$. Para eso hallaremos la solución del problema (21).

Sean $l_\alpha(x_\beta)$ y $L_\alpha(x_\beta)$ los extremos del intervalo Δ_α, sobre el cual están distribuidos los nodos de la red $\omega_\alpha(x_\beta)$. En virtud de la construcción del retículo sobre el plano los pasos $h_\alpha^\pm$ son diferentes de h sólo para los nodos fronterizos (véase la fig. 3).

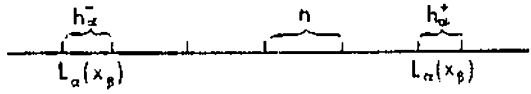

Fig. 3

Por eso sobre la red $\omega_\alpha(x_\beta)$ la derivada de diferencias $v_{\bar{x}_\alpha\hat{x}_\alpha}$ y el segundo miembro de la ecuación (21) se escriben en la forma

$$v_{\bar{x}_\alpha\hat{x}_\alpha}=\frac{1}{h}\left(\frac{v^{+1_\alpha}-v}{h}-\frac{v-v^{-1_\alpha}}{h_\alpha^-}\right),$$

$$b_\alpha=\frac{1}{h}\left(\frac{1}{h}+\frac{1}{h_\alpha^-}\right),\quad x_\alpha=l_\alpha+h_\alpha^-,$$

$$v_{\bar{x}_\alpha\hat{x}_\alpha}=\frac{1}{h^2}(v^{+1_\alpha}-2v+v^{-1_\alpha}),$$

$$b_\alpha=\frac{2}{h^2},\quad l_\alpha+h_\alpha^-<x_\alpha<L_\alpha-h_\alpha^+,$$

$$v_{\bar{x}_\alpha \hat{x}_\alpha} = \frac{1}{h}\left(\frac{v^{+1_\alpha} - v}{h_\alpha^+} - \frac{v - v^{-1_\alpha}}{h}\right),$$

$$b_\alpha = \frac{1}{h}\left(\frac{1}{h} + \frac{1}{h_\alpha^+}\right), \quad x_\alpha = L_\alpha - h_\alpha^+.$$

Una verificación directa muestra, que la función reticular

$$v^\alpha(x) = \frac{1}{h^2}\left[(x_\alpha - l_\alpha)(L_\alpha - x_\alpha + \frac{(h_\alpha^-)^2 - (h_\alpha^+)^2}{L_\alpha - l_\alpha}\right) + h^2 - (h_\alpha^-)^2\Big]$$

para $x_\alpha \in \omega_\alpha(x_\beta)$ es la solución del problema (21). Ya que

$$v^\alpha(x) \leqslant \frac{1}{h^2}(x_\alpha - l_\alpha)(L_\alpha - x_\alpha) + 1,$$

entonces

$$\varkappa_\alpha = \max_{x_\alpha \in \omega_\alpha(x_\beta)} v^\alpha(x) \leqslant \frac{1}{h^2}\left(\frac{L_\alpha - l_\alpha}{2}\right)^2 + 1. \tag{24}$$

Sustituyendo (18) y (24) en (23) hallaremos la estimación de γ_1.

Una estimación aproximada para γ_1 se puede obtener de la siguiente forma. Sea la región $\overline{G}$ examinada inscrita en un cuadrado con lado l. Entonces $L_\alpha - l_\alpha \leqslant l$ para cualquier α y, por consiguiente, $\varkappa_\alpha \leqslant l^2/(4h^2) + 1$, $\alpha = 1, 2$. Sustituyendo esta estimación en (23), obtenemos, $\gamma_1 \geqslant 4h^2/(l^2 + 4h^2)$, es decir $\gamma_1 \approx 4h^2/l^2$. Como $\gamma_2 = 2 - \gamma_1$, entonces $\xi = \gamma_1/\gamma_2 \approx 2h^2/l^2$. Por lo tanto, de la estimación (7) para el número de iteraciones obtenemos

$$n_0(\varepsilon) \approx \frac{l^2}{2\sqrt{2h}} \ln\frac{2}{\varepsilon} \approx 0{,}35N \ln\frac{2}{\varepsilon}, \tag{25}$$

donde N es el número máximo de nodos por cada dirección.

De esta forma, para el método de Chebishev implícito examinado aquí el número de iteraciones depende únicamente del paso fundamental h de la red y no depende de los pasos no uniformes $h_\alpha^\pm$ en los nodos fronterizos. Comparando la estimación (25) con la estimación (10) obtenida anteriormente, encontramos, que el número de iteraciones para el caso de una región $\overline{G}$ arbitraria, inscrita en un cuadrado con lado l, es el mismo que para el caso, cuando la región $\overline{G}$ es el cuadrado indicado.

Citemos las fórmulas de cálculo para el método iterativo de Chebishev de resolución del problema de Dirichlet de

diferencias para la ecuación de Poisson en una región $\overline{G}$ arbitraria:

$$y_{k+1}(x_{ij}) = (1-\tau_{k+1})\, y_k(x_{ij}) + \\ + \frac{\tau_{k+1}}{b_1(x_{ij})+b_2(x_{ij})}\left[\frac{1}{h}\left(\frac{y^{+1_1}}{h_1^+}+\frac{y^{-1_1}}{h_1^-}+\frac{y^{+1_2}}{h_2^+}+\right.\right. \\ \left.\left.+\frac{y^{-1_2}}{h_2^-}\right)+\varphi\right]_{x=x_{ij}},$$

$x_{ij} \in \omega, \quad y_k(x) = g(x), \quad x \in \gamma.$

Notemos, que en el capítulo X será construido para el problema examinado otro método de Chebishev implícito (método iterativo alternativo triangular), para el cual el número de operaciones aritméticas, gastadas en la realización de un paso iterativo, es un poco mayor que para el método aquí examinado, y el número de iteraciones es considerablemente menor, lo cual asegura la efectividad del método indicado.

3. Problema de Dirichlet de diferencias para una ecuación elíptica con coeficientes variables.

3.1 MÉTODO DE CHEBISHEV EXPLÍCITO. Examinemos el problema de Dirichlet para la ecuación elíptica de segundo orden con coeficientes variables en el rectángulo $G = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$:

$$Lu = \sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha}\left(k_\alpha(x)\frac{\partial u}{\partial x_\alpha}\right) - q(x)\,u = -f(x),$$

$$x \in G, \quad u(x) = g(x), \quad x \in \Gamma. \tag{26}$$

Sobre la red rectangular

$$\overline{\omega} = \{x_{ij} = (ih_1, jh_2) \in \overline{G}, \quad 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2, \\ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1, 2\}$$

al problema diferencial (26) le corresponde el problema de diferencias

$$\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha(x)\, y_{\overline{x}_\alpha})_{x_\alpha} - d(x)\,y = -\varphi(x), \quad x \in \omega, \tag{27}$$

$$y(x) = g(x), \quad x \in \gamma.$$

Si los coeficientes $k_\alpha(x)$, $q(x)$ y $f(x)$ son los funciones suficientemente suaves, entonces los coeficientes $a_\alpha(x)$

$d(x)$ y $\varphi(x)$ del esquema de diferencias (27) pueden, por ejemplo, determinarse de la siguiente forma:

$$a_1(x_{ij}) = k_1((i-0{,}5)h_1, jh_2), \quad a_2(x_{ij}) = k_2(ih_1, (j-0{,}5)h_2),$$

$$d(x_{ij}) = q(x_{ij}), \quad \varphi(x_{ij}) = f(x_{ij}).$$

Supondremos que los coeficientes del esquema de diferencias (27) satisfacen las condiciones

$$0 < c_1 \leqslant a_\alpha(x) \leqslant c_2, \quad x \in \overline{\omega},$$
$$0 \leqslant d_1 \leqslant d(x) \leqslant d_2, \quad x \in \omega, \ \alpha = 1, 2. \tag{28}$$

Estas condiciones aseguran la existencia y unicidad de solución del problema (27).

El esquema de diferencias (27) se reduce a la ecuación operacional (1) de la forma usual: $Ay = -\mathring{\Lambda}y$, donde $y \in H$, $\mathring{y} \in \mathring{H}$, y H es el espacio de las funciones reticulares definidas sobre ω, con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x) v(x) h_1 h_2,$$

y el segundo miembro f de la ecuación (1) se diferencia del miembro derecho φ del esquema (27) sólo en los nodos fronterizos.

Para la resolución aproximada de la ecuación (1) aplicaremos el método de Chebishev explícito (5)-(7) ($B = E$). En el § 2 del cap. V fue mostrado, que el operador A aquí definido es autoconjugado en H. Por eso la información a priori para el método de Chebishev tiene la forma de las constantes γ_1 y γ_2 de las desigualdades $\gamma_1 E \leqslant A \leqslant \gamma_2 E$, $\gamma_1 > 0$, si en calidad de D se toma uno de los operadores E, A o A^*. Hallemos estas constantes. Para esto introduzcamos el operador A correspondiente al operador de diferencias $\mathring{\Lambda}$, donde $\mathring{\Lambda}y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2}$ y definamos los siguientes productos escalares para las funciones reticulares, definidas sobre $\overline{\omega}$:

$$(u, v)_{\omega_\alpha} = \sum_{x_\alpha \in \omega_\alpha} u(x) v(x) h_\alpha, \quad (u, v)_{\omega_\alpha^+} = \sum_{x_\alpha \in \omega_\alpha^+} u(x) v(x) h_\alpha,$$

$$(u, v)_\alpha = ((u, v)_{\omega_\alpha^+}, 1)_{\omega_\beta}, \quad \beta = 3 - \alpha, \ \alpha = 1, 2.$$

Aquí

$$\omega_\alpha = \{x_{\alpha, i} = ih_\alpha, \quad 1 \leqslant i \leqslant N_\alpha - 1\}; \quad \omega_\alpha^+ = \\ = \{x_{\alpha, i} = ih_\alpha, \quad 1 \leqslant i \leqslant N_\alpha\}.$$

Los productos escalares, aquí introducidos, son el análogo de los productos escalares definidos en el punto 2.

De la definición de los operadores A y $\mathring{A}$ y de las fórmulas de Green de diferencias (véase el punto 2 del § 2 cap. V) obtenemos

$$(Au, u) = -(\Lambda \mathring{u}, \mathring{u}) = -((a_1 \mathring{u}_{\bar{x}_1 x_1}, \mathring{u})_{\omega_1}, 1)_{\omega_2} - \\ - ((a_2 \mathring{u}_{\bar{x}_2 x_2}, \mathring{u})_{\omega_2}, 1)_{\omega_1} + (du, u) = \sum_{\alpha=1}^{2} (a_\alpha \mathring{u}^2_{\bar{x}_\alpha}, 1)_\alpha + (du, u), \tag{29}$$

$$(\mathring{A}u, u) = -(\Lambda \mathring{u}, \mathring{u}) = \sum_{\alpha=1}^{2} (\mathring{u}^2_{\bar{x}_\alpha}, 1)_\alpha, \quad u \in H, \quad \mathring{u} \in \mathring{H}.$$

Teniendo en cuenta las desigualdades (28), de aquí obtendremos las desigualdades operacionales del tipo

$$c_1 \mathring{A} + d_1 E \leqslant A \leqslant c_2 \mathring{A} + d_2 E. \tag{30}$$

En el punto 1 del § 5 fue mostrado que el operador $\mathring{A}$ tiene cotas

$$\mathring{\gamma}_1 = \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha}, \quad \mathring{\gamma}_2 = \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2l_\alpha},$$

es decir, tienen lugar las desigualdades

$$\mathring{\gamma}_1 E \leqslant \mathring{A} \leqslant \mathring{\gamma}_2 E. \tag{31}$$

De (30) y (31) hallamos, que el operador A tiene las cotas $\gamma_1 = c_1 \mathring{\gamma}_1 + d_1$ y $\gamma_2 = c_2 \mathring{\gamma}_2 + d_2$.

De esta forma, hemos hallado las constantes γ_1 y γ_2. Aplicándolas, calculemos mediante las fórmulas (6) los parámetros iterativos τ_k y por las fórmulas (7) hallemos una estimación para el número de iteraciones n.

Ya que $\mathring{\xi} = \mathring{\gamma}_1 / \mathring{\gamma}_2 = O(|h|^2)$, entonces $\xi = \gamma_1/\gamma_2 = O(|h|^2)$ y para el número de iteraciones del método examinado tiene lugar la siguiente estimación asintótica:

$$n_0(\varepsilon) = O\left(\frac{1}{|h|} \ln \frac{2}{\varepsilon}\right),$$

y la constante en la estimación depende de las características extremales de los coeficientes $a_\alpha(x)$ y $d(x)$, es decir de c_α y d_α, $\alpha = 1, 2$.

En el caso particular, cuando $\bar{G}$ es un cuadrado con lado l ($l_1 = l_2 = l$), la red $\bar{\omega}$ es cuadrada ($h_1 = h_2 = h = l/N$) y $d \equiv 0$, obtenemos

$$\gamma_1 = \frac{8c_1}{h^2} \operatorname{sen}^2 \frac{\pi h}{2l},$$

$$\gamma_2 = \frac{8c_2}{h^2} \cos^2 \frac{\pi h}{2l}, \quad \xi = \frac{c_1}{c_2} \operatorname{tg}^2 \frac{\pi h}{2l}$$

y, por consiguiente,

$$n_0(\varepsilon) \approx \sqrt{\frac{c_2}{c_1}} \frac{l}{\pi h} \ln \frac{2}{\varepsilon} \approx 0{,}32 \sqrt{\frac{c_2}{c_1}} N \ln \frac{2}{\varepsilon}.$$

Comparando la estimación obtenida aquí para el número de iteraciones del método explícito de Chebishev de resolución de la ecuación en diferencias (27) con coeficientes variables con la estimación (10), encontramos que para el ejemplo examinado el número de iteraciones es $\sqrt{c_2/c_1}$ veces mayor que el número de iteraciones para el caso de coeficientes constantes.

Citemos las fórmulas de cálculo para el método de Chebishev explícito (5)-(7) utilizado en la resolución de la ecuación en diferencias (27). Estas fórmulas tienen la forma:

$$y_{k+1}(i, j) = \alpha_{k+1}(i, j) y_k(i, j) + \\ + \tau_{k+1} \left\{ \frac{1}{h_1^2} [a_1(i+1, j) y_k(i+1, j) + \right. \\ + a_1(i, j) y_k(i-1, j)] + \frac{1}{h_2^2} [a_2(i, j+1) y_k(i, j+1) + \\ \left. + a_2(i, j) y_k(i, j-1)] + \varphi(i, j) \right\},$$

$1 \leqslant i \leqslant N_1 - 1$, $1 \leqslant j \leqslant N_2 - 1$,

donde está designado

$$\alpha_{k+1}(i, j) = 1 - \tau_{k+1} \left[\frac{a_1(i+1, j) + a_1(i, j)}{h_1^2} + \right. \\ \left. + \frac{a_2(i, j+1) + a_2(i, j)}{h_2^2} + d(i, j) \right],$$

y la aproximación inicial y_0 es una función reticular arbitraria sobre ω, que toma sobre γ los valores prefijados: $y(x) = g(x)$ para $x \in \gamma$.

3.2 МЕТОДО DE CHEBISHEV IMPLÍCITO. Para la resolución aproximada de la ecuación (1) planteada en el punto anterior, correspondiente al esquema de diferencias (27), examinemos ahora el método de Chebishev implícito más simple (5)-(7). En calidad de operador B, igual que en el punto 2, de nuevo tomamos la parte diagonal del operador A

$$By = by,$$

$$b(i, j) = \frac{1}{h_1^2}[a_1(i+1, j) + a_1(i, j)] + \\ + \frac{1}{h_2^2}[a_2(i, j+1) + a_2(i, j)] + d(i, j). \qquad (32)$$

Puesto que los operadores A y B son autoconjugados y definidos positivos en H, entonces la información a priori para el método de Chebishev implícito (5)-(7) tiene la forma de las constantes de equivalencia energética de los operadores $\gamma_1 B \leqslant A \leqslant \gamma_2 B$, $\gamma_1 > 0$, si en calidad de D es elegido uno de los operadores A, B ó $AB^{-1}A$.

Hallemos las constantes γ_1 y γ_2. Al igual que en el punto 2, se demuestra que tiene lugar la igualdad $\gamma_1 + \gamma_2 = 2$. Por eso en las fórmulas (6) para los parámetros iterativos τ_k tenemos $\tau_0 = 2/(\gamma_1 + \gamma_2) = 1$, $\rho_0 = (\gamma_2 - \gamma_1)/\gamma_2 + \gamma_1) = 1 - \gamma_1$.

Estimemos γ_1. Del lema 14 del § 2 cap. V se deduce que para cualquier función reticular $\mathring{y} \in \mathring{H}$ tiene lugar la desigualdad

$$(b\mathring{y}, \mathring{y})_{\omega_\alpha} \leqslant \varkappa_\alpha \left[\left(a_\alpha \mathring{y}^2_{\bar{x}_\alpha}, 1\right)_{\omega_\alpha^+} + \frac{1}{2}(d\mathring{y}, \mathring{y})_{\omega_\alpha} \right],$$

$$\alpha = 1, 2, \qquad (33)$$

donde $\varkappa_\alpha = \varkappa_\alpha(x_\beta) = \max\limits_{x_\alpha \in \omega_\alpha} v^\alpha(x)$, y $v^\alpha(x)$ es la solución del siguiente problema de contorno tripuntual:

$$(a_\alpha v^\alpha_{\bar{x}_\alpha})_{x_\alpha} - \frac{1}{2} dv^\alpha = -b(x), \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,$$

$$v^\alpha(x) = 0, \; x_\alpha = 0, \; l_\alpha, \; h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta, \; \beta = 3 - \alpha, \qquad (34)$$

$$\alpha = 1, 2.$$

De las desigualdades (33) dividiendo por $\varkappa_\alpha$ y con la subsiguiente sumación según ω_β obtendremos

$$\left(\frac{b}{\varkappa_\alpha}\mathring{y}, \mathring{y}\right) \leqslant \left(a_\alpha \mathring{y}^2_{\bar{x}_\alpha}, 1\right) + \frac{1}{2}(d\mathring{y}, \mathring{y}), \quad \alpha = 1, 2.$$

Sumando estas desigualdades y teniendo en cuenta (29), obtenemos

$$\Big(\sum_{\alpha=1}^{2}\frac{1}{\varkappa_\alpha}\overset{\circ}{by},\ \overset{\circ}{y}\Big)\leqslant\sum_{\alpha=1}^{2}\Big(a_\alpha\overset{\circ}{y}^2_{x_\alpha},\ 1\Big)+(\overset{\circ}{dy},\ \overset{\circ}{y})=(Ay,\ y).$$

Por consiguiente, en calidad de γ_1 se puede tomar

$$\gamma_1=\min_{x\in\omega}\sum_{\alpha=1}^{2}\frac{1}{\varkappa_\alpha}=\min_{x_2\in\omega_2}\frac{1}{\varkappa_1(x_2)}+\min_{x_1\in\omega_1}\frac{1}{\varkappa_2(x_1)}. \qquad (35)$$

Así, para encontrar γ_1 es necesario resolver las ecuaciones (34), hallar $\varkappa_\alpha(x_\beta)$ y por la fórmula (35) calcular γ_1. La constante γ_2 se encuentra mediante la fórmula $\gamma_2 = 2-\gamma_1$.

Obtengamos una estimación para el número de iteraciones del método de Chebishev implícito a examinar. De la teoría de los esquemas de diferencias se desprende que el esquema de diferencias (34) es estable respecto al segundo miembro en la métrica uniforme, es decir, existe una constante M, no independiente de los pasos h_1 y h_2 de la red, tal que para la solución de la ecuación (34) tiene lugar la estimación

$$\max_{x_\alpha\in\omega_\alpha} v^\alpha(x)\leqslant M\,(b^2,\ 1)^{\frac{1}{2}}_{\omega_\alpha}.$$

Como $b(x)=O\left(\frac{1}{h^2}\right)$, $h=\min\limits_{\alpha}h_\alpha$ para $x\in\omega$, entonces de aquí obtenemos, que

$$\varkappa_\alpha=\max_{x_\alpha\in\omega_\alpha}v^\alpha(x)=O\left(\frac{1}{h^2}\right)$$

y, por consiguiente, $\gamma_1=O(h^2)$ y $\gamma_2=O(1)$. Por eso $\xi=\gamma_1/\gamma_2=O(h^2)$, y para el número de iteraciones tiene lugar la misma estimación asintótica por h que para el método explícito

$$n_0(\varepsilon)=O\left(\frac{1}{h}\ln\frac{2}{\varepsilon}\right).$$

¿En que consiste la ventaja del método iterativo implícito en comparación con el método de Chebishev explícito examinado más arriba? La respuesta a esta pregunta la da el siguiente teorema, el cual citaremos sin demostración.

TEOREMA 7 *. *Para el esquema iterativo (5)-(7) con el operador A correspondiente al esquema de diferencias (27), el mejor de los operadores B en la clase de los operadores diagonales (es decir aquel para el cual el cociente* ξ *es máximo) es el operador, definido por las fórmulas (32).*

Del teorema 7 se deduce que si en calidad del operador B se elige la parte diagonal del operador A, entonces la relación $\xi = \gamma_1/\gamma_2$ de las constantes de equivalencia energética γ_1 y γ_2 de los operadores A y B será máximo, y por lo tanto, el número de iteraciones n será mínimo.

Ilustremos la ventaja del método implícito en el siguiente ejemplo. Sea prefijado el esquema de diferencias (27) sobre una red rectangular en el cuadrado unidad $h_1 = h_2 = h = 1/N$, $l_1 = l_2 = 1$.

Eligiremos los coeficientes $a_1(x)$, $a_2(x)$ y $d(x)$ de la siguiente manera:

$$a_1(x) = 1 + c\,[(x_1 - 0{,}5)^2 + (x_2 - 0{,}5)^2],$$
$$a_2(x) = 1 + c\,[0{,}5 - (x_1 - 0{,}5)^2 - (x_2 - 0{,}5)^2],$$
$$d(x) \equiv 0,\ c > 0.$$

Con esto en las desigualdades (28) tenemos $c_1 = 1$, $c_2 = 1 + 0{,}5c$, $d_1 = d_2 = 0$. Cambiando el parámetro c, obtendremos los coeficientes del esquema de diferencias (27) con características extremales diferentes.

Para el método explícito fue mostrado, que el número de iteraciones depende de la relación c_2/c_1. Para el método implícito el número de iteraciones no depende de los valores máximo y mínimo de los coeficientes $a_\alpha(x)$, sino de algunas características integrales de estos coeficientes.

Tabla 7

$\frac{c_2}{c_1}$	$N=32$		$N=64$		$N=128$	
	implícito	explícito	implícito	explícito	implícito	explícito
2	123	143	246	286	494	571
8	149	286	305	571	616	1142
32	175	571	365	1142	749	2283
128	192	1141	409	2283	856	4565
512	202	2281	436	4565	926	9130

*) Este teorema es un caso particular de un teorema más general demostrado en el trabajo: G. Forsythe, E. G. Straus, On best conditioned matrices, Proc. Amer. Math. Soc. 6 (1955), 340-345.

En la tabla 7 se expone el número de iteraciones para los métodos explícito e implícito en función de la relación c_2/c_1 y del número de nodos N por una dirección. Los cálculos están hechos para $\varepsilon = 10^{-4}$. Para el caso, cuando el parámetro $c = 0$, es decir, $a_\alpha(x) \equiv 1$, y se examina la ecuación de Poisson, el número de iteraciones para los métodos explícito e implícito es el mismo y se da en el punto 1.

De la tabla se deduce, que para el ejemplo a examinar el número de iteraciones para el método implícito es considerablemente menor, que para el método explícito. La dependencia del número de iteraciones de la relación c_2/c_1 es menor para el método implícito que para el explícito.

En resumen citemos las fórmulas de cálculo para el método Chebishev implícito:

$$y_{k+1}(i, j) = (1 - \tau_{k+1}) y_k(i, j) + \\ + \frac{\tau_{k+1}}{b(i, j)} \left\{ \frac{1}{h_1^2} [a_1(i+1, j) y_k(i+1, j) + \right. \\ + a_1(i, j) y_k(i-1, j)] + \frac{1}{h_2^2} [a_2(i, j+1) y_k(i, j+1) + \\ \left. + a_2(i, j) y_k(i, j+1)] + \varphi(i, j) \right\},$$

$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1,$$

donde $b(i, j)$ está definido en (32), y la aproximación inicial y_0 es una función reticular arbitraria sobre ω, que toma sobre la frontera γ los valores prefijados: $y_0(x) = g(x)$, $x \in \gamma$.

De la comparación de las fórmulas de cálculo para los métodos de Chebishev explícito e implícito se deduce, que el número de operaciones aritméticas para calcular y_{k+1} por el y_k dado es prácticamente igual para ambos métodos. Puesto que el número de iteraciones para el método implícito es considerablemente menor, que para el explícito, entonces se le debe dar preferencia al método implícito.

4. Problema de Dirichlet de diferencias para una ecuación elíptica con derivada mixta. En el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \ \alpha = 1, 2\}$ con frontera Γ se exige resolver *el problema de Dirichlet para una ecuación elíptica con derivadas mixtas*

$$Lu = \sum_{\alpha, \beta = 1}^{2} \frac{\partial}{\partial x_\alpha} \left(k_{\alpha, \beta}(x) \frac{\partial u}{\partial x_\beta} \right) = -f(x), \quad x \in G,$$

$$u(x) = g(x), \quad x \in \Gamma.$$

Se supone, que se cumplen las condiciones de simetría

$$k_{12}(x) = k_{21}(x),\ x \in \overline{G}, \tag{36}$$

y de elipticidad

$$c_1 \sum_{\alpha=1}^{2} \xi_\alpha^2 \leqslant \sum_{\alpha,\,\beta=1}^{2} k_{\alpha\beta}\xi_\alpha\xi_\beta \leqslant c_2 \sum_{\alpha=1}^{2} \xi_\alpha^2,\ c_1 > 0, \tag{37}$$

donde $\xi = (\xi_1, \xi_2)$ es un vector arbitrario.

Sobre la red rectangular

$$\overline{\omega} = \{x_{ij} = (ih_1,\ jh_2 \in \overline{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,$$
$$h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1,\ 2\}$$

al problema diferencial le corresponde el problema de Dirichlet de diferencias

$$\Lambda y = 0{,}5 \sum_{\alpha,\,\beta=1}^{2} [(k_{\alpha\beta} y_{\overline{x}_\beta})\, x_\alpha + (k_{\alpha\beta} y_{x_\beta})_{\overline{x}_\alpha}] = -\varphi(x),\ x \in \omega, \tag{38}$$

$$y(x) = g(x),\ x \in \gamma,$$

donde γ es la frontera de la red $\overline{\omega}$.

En el § 2 del cap. V fue mostrado, que el problema de diferencias (38) se reduce a la ecuación operacional (1) de la manera usual: $Ay = -\Lambda\mathring{y}$, donde $y \in H$, $\mathring{y} \in \mathring{H}$ e $\mathring{y}(x) = y(x)$ para $x \in \omega$. Aquí $\mathring{H}$ es el conjunto de funciones reticulares, definidas sobre $\overline{\omega}$ y que se anulan sobre γ, y H es el espacio de las funciones reticulares, definidas sobre ω, con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2.$$

Allí también fue mostrado, que al satisfacer las condiciones (36) el operador A construido es autoconjugado en H y, si se cumplen las condiciones (37), él posee las cotas γ_1

y γ_2, iguales a

$$\gamma_1 = c_1 \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha}, \quad \gamma_2 = c_2 \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2l_\alpha}, \tag{39}$$

es decir

$$\gamma_1 E \leqslant A \leqslant \gamma_2 E. \tag{40}$$

Para la resolución aproximada de la ecuación (1), correspondiente al esquema de diferencias (38), examinaremos el método de Chebishev explícito (5)-(7) ($B = E$). Ya que los operadores A y B son autoconjugados y definidos positivos en H, entonces la información a priori tiene la forma de las constantes γ_1 y γ_2 en las desigualdades (40), y el método converge en H_D, donde $D = A$, B ó $AB^{-1}A$.

De (39) obtenemos

$$\gamma_1 = O(c_1), \quad \gamma_2 = O\left(\frac{c_2}{h^2}\right), \quad \xi = \frac{\gamma_1}{\gamma_2} = O\left(\frac{c_1}{c_2} h^2\right),$$

$$h^2 = h_1^2 + h_2^2.$$

Por lo tanto, para examinar el ejemplo, la estimación asintótica respecto de h del número de iteraciones $n_0(\varepsilon)$ tiene la forma

$$n_0(\varepsilon) = O\left(\sqrt{\frac{c_2}{c_1}} \frac{1}{h} \ln \frac{2}{\varepsilon}\right).$$

En el caso particular, cuando $\overline{G}$ es un cuadrado con lado l y la red $\overline{\omega}$ es cuadrado ($h_1 = h_2 = h \quad l/N$), obtenemos

$$\gamma_1 = \frac{8c_1}{h^2} \operatorname{sen}^2 \frac{\pi h}{2l},$$

$$\gamma_2 = \frac{8c_2}{h^2} \cos^2 \frac{\pi h}{2l}, \quad \xi = \frac{c_1}{c_2} \operatorname{tg}^2 \frac{\pi h}{2l},$$

$$n \geqslant n_0(\varepsilon) \approx \sqrt{\frac{c_2}{c_1}} \frac{l}{\pi h} \ln \frac{2}{\varepsilon} = 0{,}32 \sqrt{\frac{c_2}{c_1}} N \ln \frac{2}{\varepsilon},$$

es decir, el número de iteraciones es igualmente proporcional al número N de nodos por una dirección, como también en el caso de la ecuación sin derivadas mixtas.

Con esto nosotros terminamos el examen de los ejemplos de aplicación de los métodos iterativos de dos capas a la resolución de ecuaciones elípticas. Ejemplos más complejos serán examinados en el capítulo XIV.

Capítulo VII

Métodos iterativos de tres capas

En este capítulo se estudian los métodos iterativos de tres capas para resolver la ecuación operacional $Au = f$. Los parámetros iterativos se eligen teniendo en cuenta la información a priori sobre los operadores del esquema. En el § 1 se da la estimación de la velocidad de convergencia para los esquemas de tres capas de tipo estándar. En los §§ 2 y 3 son examinados el método semi-iterativo de Chebishev y el método de tres capas estacionario. El § 4 está dedicado a la investigación de la estabilidad de los métodos de dos y tres capas con respecto a la perturbación de los datos.

§ 1. Estimación de la velocidad de convergencia

1. Familia inicial de esquemas iterativos. En el capítulo VI para encontrar una solución aproximada de la ecuación operacional lineal

$$Au = f \tag{1}$$

con el operador no degenerado A, el cual actúa en un espacio de Hilbert real H, fueron construidos los métodos iterativos de dos capas. En estos métodos el esquema de dos capas relaciona dos aproximaciones iterativas y_{k+1} y y_k.

En este capítulo serán estudiados los esquemas iterativos de tres capas. El esquema iterativo de tres capas para la ecuación (1) relaciona tres aproximaciones iterativas y_{k+1}, y_k e y_{k-1}, de manera tal que y_{k+1} se determina por medio de y_k e y_{k-1}. Para la realización de un esquema de tres capas deben ser prefijadas dos aproximaciones iniciales y_0 e y_1. Con frecuencia para un y_0 arbitrario la aproximación y_1 se encuentra por un esquema de dos capas.

Limitémonos al estudio de los esquemas de tres capas de tipo estandar. *Un esquema iterativo de tres capas estandar* implícito tiene la forma

$$By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)\, y_k + (1 - \alpha_{k+1})\, By_{k-1} + \\ + \alpha_{k+1}\tau_{k+1}f, \qquad (2)$$
$$By_1 = (B - \tau_1 A)\, y_0 + \tau_1 f, \quad k = 1, 2, \ldots, y_0 \in H,$$

donde y_0, es una aproximación inicial arbitraria, B es un operador lineal no degenerado que actúa en H, α_k y τ_k son los parámetros iterativos. Por las fórmulas (2) se define una familia inicial de esquemas iterativos de tres capas.

El encuentro de una nueva aproximación y_{k+1} se puede interpretar de modo siguiente. Sea $\bar{y}$ una aproximación iterativa intermedia, la cual se encuentra por el esquema implícito de dos capas

$$B\frac{\bar{y} - y_k}{\tau_{k+1}} + Ay_k = f.$$

Entonces de (2) se deduce, que y_{k+1} es una combinación lineal de las aproximaciones $\bar{y}$ e y_{k-1}

$$y_{k+1} = \alpha_{k+1}\bar{y} + (1 - \alpha_{k+1})\, y_{k-1}.$$

De esta forma, la aproximación y_{k+1} es una extrapolación lineal según las aproximaciones $\bar{y}$ e y_{k-1}.

Si ponemos $\alpha_k \equiv 1$ en (2), entonces el esquema de tres capas pasa a ser un esquema de dos capas cuya convergencia fue estudiada en el capítulo VI. Por eso la introducción de los parámetros iterativos α_k permite presuponer que la convergencia del esquema (2) no será peor que la convergencia de un esquema de dos capas.

Señalemos, que a diferencia de un esquema iterativo de dos capas, la realización de un esquema de tres capas exige la recordación no solamente de una, sino de dos aproximaciones iterativas y_k e y_{k-1}.

2. Estimación de la norma del error. Ocupémonos ahora de la investigación de la convergencia del esquema de tres capas (2) en el espacio energético H_D, generado por un operador D autoconjugado y definido positivo en H. Para eso estudiaremos el comportamiento de la norma del error $z_k = y_k - u$ cuando $k \to \infty$ en H_D.

Sustituyendo $y_k = z_k + u$ para $k = 0, 1, \ldots$ en (2) y teniendo en cuenta la ecuación (1), hallaremos una ecuación para el error z_k:

$$Bz_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)\, z_k + (1 - \alpha_{k+1})\, Bz_{k-1},$$
$$k = 1, 2, \ldots,$$
$$Bz_1 = (B - \tau_1 A)\, z_0, \quad z_0 = y_0 - u.$$

Resolvamos esta ecuación con respecto a z_{k+1} y, poniendo $z_k = D^{-1/2}x_k$, pasemos a la ecuación para el error equivalente x_k. La ecuación para x_k tendrá la forma siguiente:

$$x_{k+1} = \alpha_{k+1}S_{k+1}x_k + (1 - \alpha_{k+1})\, x_{k-1}, \quad k = 1, 2, \ldots,$$
$$x_1 = S_1 x_0, \quad S_k = E - \tau_k C, \tag{3}$$

donde $C = D^{1/2}B^{-1}AD^{-1/2}$.

En virtud del cambio hecho $z_k = D^{-1/2}x_k$ es válida la igualdad $\| x_k \| = \| z_k \|_D$ y, por consiguiente, la convergencia del esquema (2) en H_D tendrá lugar, si $\| x_k \| \to 0$ cuando $k \to \infty$.

Estudiemos el comportamiento de la norma de x_k en H cuando $k \to \infty$. Para eso hallemos la forma explícita de la solución de la ecuación (9). Utilizando las fórmulas (9), sucesivamente obtendremos

$$x_1 = (E - \tau_1 C)\, x_0 = P_1(C)\, x_0,$$
$$x_2 = \alpha_2 (E - \tau_2 C)\, x_1 + (1 - \alpha_2)\, x_0 =$$
$$= [\alpha_2 (E - \tau_2 C)\, P_1(C) + (1 - \alpha_2)\, E]\, x_0 = P_2(C)\, x_0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x_{k+1} = \alpha_{k+1}(E - \tau_{k+1}C)\, P_k(C)\, x_0 +$$
$$+ (1 - \alpha_{k+1})\, P_{k-1}(C)\, x_0 = P_{k+1}(C)\, x_0,$$

etc.

Por lo tanto, para cada k la solución de la ecuación (9) tiene la forma

$$x_k = P_k(C)\, x_0, \quad k = 0, 1, \ldots, \tag{4}$$

donde $P_k(C)$ es el polinomio operacional de grado k respecto al operador C. En virtud de la arbitrariedad de x_0, el correspondiente polinomio algebraico $P_k(t)$ satisface las siguientes relaciones recurrentes:

$$P_{k+1}(t) = \alpha_{k+1}(1 - \tau_{k+1}t)\, P_k(t) + (1 - \alpha_{k+1})P_{k-1}(t),$$
$$P_1(t) = 1 - \tau_1 t, \quad P_0(t) \equiv 1, \quad k = 1, 2, \ldots \tag{5}$$

De (5) se deduce que para cualquier k el polinomio $P_k(t)$ satisface la condición de normación $P_k(0) = 1$.

Estimemos ahora la norma de x_k. De (4) obtenemos

$$\| x_k \| = \| P_k(C) x_0 \| \leqslant \| P_k(C) \| \| x_0 \|,$$
$$k = 0, 1, \ldots,$$

o, en virtud de la sustitución hecha $z_k = D^{-\frac{1}{2}} x_k$,

$$\| z_k \|_D \leqslant \| P_k(C) \| \| z_0 \|_D. \tag{6}$$

Así, está obtenida la estimación de la norma del error z_k. De (6) se desprende que el método poseerá la mayor velocidad de convergencia, si la norma del polinomio $P_k(C)$ tiende a cero cuando $k \to \infty$ lo más rápido posible. Puesto que el polinomio $P_k(C)$ es una función de los parámetros iterativos $\tau_1, \tau_2, \ldots, \tau_k$ y $\alpha_2, \alpha_3, \ldots, \alpha_k$, entonces estos parámetros deben ser elegidos de la condición de mínimo de la norma del polinomio operacional $P_k(C)$. En otras palabras, es necesario construir un polinomio de grado k normado por la condición $P_k(0) = E$, el cual posea norma mínima.

En el capítulo VI durante el estudio del método de Chebishev de dos capas, este problema fue resuelto suponiendo que el operador $DB^{-1}A$ es autoconjugado en H y siendo prefijadas las constantes de equivalencia energética γ_1 y γ_2 de los operadores autoconjugados D y $DB^{-1}A$:

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0, \quad DB^{-1}A = (DB^{-1}A)^*. \tag{7}$$

Al construir los métodos iterativos de tres capas vamos a examinar únicamente el caso autoconjugado, es decir, supondremos que están cumplidas las condiciones (7).

Dado que están cumplidas las condiciones (7). Indicaremos el operador óptimo $P_k(C)$ y obtendremos la estimación a priori para el error z_k.

Ya que $C = D^{\frac{1}{2}} B^{-1} A D^{-\frac{1}{2}} = D^{-\frac{1}{2}} (DB^{-1}A) D^{-\frac{1}{2}}$, entonces de (7) se deduce, que el operador C es autoconjugado en H, y γ_1, γ_2 son sus cotas:

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \gamma_1 > 0, \quad C = C^*. \tag{8}$$

Entonces en virtud de (8) para la norma del operador $P_k(C)$, es válida la estimación

$$\| P_k(C) \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} | P_k(t) |.$$

Por consiguiente, el polinomio óptimo $P_k(t)$ se destaca por la siguiente condición: que el máximo del módulo de este polinomio sobre el segmento $[\gamma_1, \gamma_2]$ sea mínimo. Del § 2 del cap. VI se deduce, que para la condición de normación $P_k(0) = 1$ el polinomio buscado tiene la forma

$$P_k(t) = q_k T_k\left(\frac{1-\tau_0 t}{\rho_0}\right), \quad q_k = \frac{1}{T_k\left(\frac{1}{\rho_0}\right)}, \tag{9}$$

donde $T_k(x)$ es el polinomio de Chebishev de 1 er. género y de grado k:

$$T_k(x) = \begin{cases} \cos(k \arccos x), & |x| \leqslant 1, \\ \operatorname{ch}(k \operatorname{Arch} x), & |x| \geqslant 1, \end{cases}$$

$$\tau_0 = \frac{2}{\gamma_1+\gamma_2}, \quad \rho_0 \frac{1-\xi}{1+\xi}, \quad q_k = \frac{2\rho_1^k}{1+\rho_1^{2k}},$$

$$\rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Con esto tiene lugar la estimación

$$\| P_k(C) \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |P_k(t)| = q_k, \quad k = 0, 1, \ldots$$

Sustituyendo esta estimación en (6), obtenemos

$$\| z_k \|_D \leqslant q_k \| z_0 \|_D.$$

De esta manera, la velocidad de convergencia del método iterativo de tres capas (2), cuyos parámetros τ_k y α_k se eligen de la condición de mínimo de la norma del operador de resolución, es igual a la velocidad de convergencia del método iterativo de Chebishev de dos capas.

Las fórmulas (9) dan la solución del problema sobre la construcción del método iterativo de tres capas de convergencia más rapida. En el § 2 serán obtenidas las fórmulas para los parámetros iterativos τ_k y α_k de este método, el cual se llama *método semi-iterativo de Chebishev.*

§ 2. Método semi-iterativo de Chebishev

1. Fórmulas para los parámetros iterativos. Hallamos ahora las fórmulas para los parámetros iterativos α_k y τ_k del *método semi-iterativo de Chebishev.* En el § 1, utilizando

el esquema iterativo de tres capas del método

$$By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)\, y_k + (1 - \alpha_{k+1})\, By_{k-1} + \\ + \alpha_{k+1}\tau_{k+1}f, \qquad (1)$$

$$By_1 = (B - \tau_1 A)\, y_0 + \tau_1 f, \quad k = 1, 2, \ldots, y_0 \in H,$$

nosotros obtuvimos una ecuación para el error equivalente

$$x_{k+1} = \alpha_{k+1}(E - \tau_{k+1}C)\, x_k + (1 - \alpha_{k+1})\, x_{k-1}, \\ k = 1, 2, \ldots, \qquad (2)$$

$$x_1 = (E - \tau_1 C)\, x_0.$$

Fue mostrado, que para cualquier k la solución de esta ecuación tiene la forma

$$x_k = P_k(C)\, x_0, \quad k = 0, 1, \ldots, \qquad (3)$$

y el polinomio $P_k(C)$ se determina por las fórmulas

$$P_k(t) = q_k T_k\left(\frac{1 - \tau_0 t}{\rho_0}\right), \quad q_k = \frac{1}{T_k\left(\frac{1}{\rho_0}\right)} = \frac{2\rho_1^k}{1 + \rho_1^{2k}} \qquad (4)$$

Para obtener las fórmulas para los parámetros iterativos α_k y τ_k, hallaremos las relaciones recurrentes, a las que satisface el polinomio $P_k(t)$.

Es conocido, que para cualquier x los polinomios de Chebishev de primer género $T_k(x)$ satisfacen las siguientes relaciones recurrentes (véase el § 4 del cap. 1):

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad k = 1, 2, \ldots, \\ T_1(x) = x, \quad T_0(x) \equiv 1. \qquad (5)$$

Utilizando (4), y (5), obtendremos

$$\frac{P_{k+1}(t)}{q_{k+1}} = 2\left(\frac{1 - \tau_0 t}{\rho_0}\right)\frac{P_k(t)}{q_k} - \frac{P_{k-1}(t)}{q_{k-1}}, \quad k = 1, 2, \ldots, \qquad (6)$$

$$P_1(t)/q_1 = (1 - \tau_0 t)/\rho_0, \quad P_0(t)/q_0 \equiv 1. \qquad (7)$$

De la definición (4) y las relaciones (5) se deduce

$$1/q_{k+1} = 2/(\rho_0 q_k) - 1/q_{k-1}, \quad q_1 = \rho_0, \quad q_0 = 1. \qquad (8)$$

De aquí hallamos

Sustituyendo (8), y (9) en (6) y (7), obtendremos las fórmulas recurrentes para los polinomios $P_k(t)$:

$$P_{k+1}(t) = \frac{2}{\rho_0}\frac{q_{k+1}}{q_k}(1-\tau_0 t)\,P_k(t) + \left(1-\frac{2}{\rho_0}\frac{q_{k+1}}{q_k}\right)P_{k-1}(t),$$

$$P_1(t) = 1-\tau_0 t, \quad P_0(t) \equiv 1, \quad k = 1, 2, \ldots$$

De aquí y de (3) se deducen las relaciones recurrentes para x_k

$$x_{k+1} = \frac{2}{\rho_0}\frac{q_{k+1}}{q_k}(E-\tau_0 C)\,x_k + \left(1-\frac{2}{\rho_0}\frac{q_{k+1}}{q_k}\right)k_{k-1},$$

$$k = 1, 2, \ldots,$$

$$x_1 = (E-\tau_0 C)\,x_0.$$

Comparando estas fórmulas con (2), obtenemos

$$\alpha_{k+1} = 2q_{k+1}/(\rho_0 q_k), \quad \tau_k \equiv \tau_0 = 2/(\gamma_1+\gamma_2), \quad k = 1, 2, \ldots \tag{10}$$

Así, hemos obtenido las fórmulas para los parámetros iterativos τ_k y α_k. Transformemos la fórmula para los parámetros α_k. Para eso, utilizando (8), calculemos la expresión

$$\frac{4\alpha_{k+1}-1}{\alpha_k\alpha_{k+1}} = \frac{1}{\alpha_k}\left(1-\frac{1}{\alpha_{k+1}}\right) = \frac{4}{\alpha_k}\left(1-\frac{\rho_0}{2}\frac{q_k}{q_{k+1}}\right) =$$

$$= \frac{4}{\alpha_k}\left[1-\frac{\rho_0}{2}\left(\frac{2}{\rho_0}-\frac{q_k}{q_{k-1}}\right)\right] = \frac{2\rho_0}{\alpha_k}\frac{q_k}{q_{k-1}} = \rho_0^2.$$

De aquí obtenemos $\alpha_{k+1} = 4/(4-\rho_0^2\alpha_k)$, $k = 1, 2, \ldots$ Poniendo $k=0$ en (10) y teniendo en cuenta (8), hallaremos, que $\alpha_1 = 2$.

Así, está demostrado

TEOREMA 1. *Supongamos que se cumplen las condiciones*

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0, \quad DB^{-1}A = (DB^{-1}A)^*$$

El método semi-iterativo de Chebishev (2) con los parámetros iterativos

$$\tau_k \equiv 2/(\gamma_1+\gamma_2), \quad \alpha_{k+1} = 4/(4-\rho_0^2\alpha_k),$$

$$k = 1, 2, \ldots, \alpha_1 = 2, \tag{11}$$

converge en H_D, y para el error z_k es válida la estimación

$$\|z_k\|_D \leqslant q_k \|z_0\|_D.$$

Para el número de iteraciones n tiene lugar la estimación $n \geqslant n_0(\varepsilon)$, *donde*

$$n_0(\varepsilon) = \frac{\ln 0{,}5\varepsilon}{\ln \rho_1}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}},$$

$$q_k = \frac{2\rho_1^k}{1+\rho_1^{2k}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

OBSERVACIÓN. La comparación del método semi-iterativo de Chebishev con el método de Chebishev de dos capas muestra, que para estos métodos tiene lugar la misma estimación $\| z_n \|_D \leqslant q_n \| z_0 \|_D$, si se han realizado n iteraciones. Sin embargo, para el método de dos capas esta estimación es cierta solamente después de ser cumplimentadas todas las iteraciones, a la vez que para el método de tres capas este tipo de estimación tiene lugar para cualquier iteración intermedia. A diferencia del método de dos capas en el método de tres capas las normas de los errores en las iteraciones intermedias decrecen monótonamente y esto asegura la estabilidad computacional de dicho método.

2. **Ejemplos de elección del operador D.** Mostremos ahora ejemplos de elección del operador D y las exigencias, impuestas a los operadores A y B, para los cuales serán cumplidas las condiciones del teorema 1.

En el punto 3 del § 2 cap. VI fueron examinados algunos casos de elección del operador D en dependencia de las propiedades de los operadores A y B. Citemos estos resultados.

1) Si los operadores A y B son autoconjugados y definidos positivos en H, entonces en calidad de D se puede elegir uno de los siguientes operadores: A, B ó $AB^{-1}A$. Con esto la información a priori puede ser prefijada en la forma

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0. \tag{12}$$

2) Si los operadores A y B son autoconjugados, definidos positivos y conmutan $A = A^* > 0$, $B = B^* > 0$, $AB = BA$, entonces en calidad de D se puede tomar el operador A^*A. En este caso la información a priori tiene la forma de las desigualdades (12).

3) Si A y B son los operadores no degenerados, que satisfacen la condición $B^*A = A^*B$, entonces en calidad de D también se puede tomar el operador A^*A. En este caso

la información, a priori se prefija en forma de las desigualdades

$$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx) \leqslant \gamma_2 (Bx, Bx), \quad \gamma_1 > 0.$$

Siendo cumplidas estas suposiciones es cierto el teorema 1 para el método semi-iterativo de Chebishev.

3. **Algoritmo del método.** Examinemos el problema de la realización del esquema de tres capas (1). El algoritmo del método puede ser descrito de la forma siguiente:

1) por el valor del parámetro α_k y por las aproximaciones dadas y_{k-1} e y_k encontramos α_{k+1} y τ_{k+1} mediante las fórmulas (11) y calculamos

$$\varphi = B(\alpha_{k+1} y_k + (1 - \alpha_{k+1}) y_{k-1}) - \alpha_{k+1}\tau_{k+1}(Ay_k - f).$$

El valor de φ calculado puede ser empleado en el lugar de y_{k-1}, el cual ya no es necesario para las iteraciones siguientes;

2) para encontrar una nueva aproximación y_{k+1} resolvemos la ecuación $By_{k+1} = \varphi$. La aproximación y_1 se encuentra de la ecuación $By_1 = \varphi$, donde $\varphi = By_0 - \tau_1 (Ay_0 - f)$. Este algoritmo de resolución puede ser recomendado en el caso, cuando se exige ahorrar memoria de la MCE.

Si para calcular el valor del operador B se necesita un mayor número de operaciones aritméticas, y no hay limitaciones de memoria, entonces es más útil emplear el siguiente algoritmo:

1) Por y_k prefijado se calcula el error $r_k = Ay_k - f$;

2) se resuelve la ecuación para la corrección w_k: $Bw_k = r_k$;

3) por α_k dado se calcula α_{k+1} según la fórmula (11) y la nueva aproximación se encuentra mediante la fórmula

$$y_{k+1} = \alpha_{k+1} y_k + (1 - \alpha_{k+1}) y_{k-1} - \alpha_{k+1}\tau_{k+1} w_k,$$

donde τ_{k+1} se determina por la fórmula (11).

El algoritmo descrito no contiene el cálculo del valor del operador B, pero requiere memoria complementaria para conservar a r_k y w_k.

§ 3. Método de tres capas estacionario

1. **Elección de los parámetros iterativos.** Regresemos ahora a las fórmulas para los parámetros iterativos α_k y τ_k del método semi-iterativo de Chebishev. En el § 1 fueron

obtenidas las siguientes expresiones para α_{k+1} y τ_{k+1}:

$$\alpha_{k+1} = 2q_{k+1}/(\rho_0 q_k), \quad \tau_k \equiv \tau_0 = 2/(\gamma_1 + \gamma_2), \quad k = 1, 2, \ldots, \tag{1}$$

donde

$$q_k = \frac{2\rho_1^k}{1+\rho_1^{2k}}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}. \tag{2}$$

El valor del parámetro iterativo τ_k no depende del número de la iteración k, en tanto que el parámetro α_k cambia, comenzando por $\alpha_1 = 2$. Hallemos el valor límite para α_k, cuando k tiende al infinito. De (1) y (2) obtenemos

$$\alpha_{k+1} = 2\rho_1 (1 + \rho_1^{2k})/(\rho_0 (1 + \rho_1^{2k+2})).$$

Ya que $\rho_1 < 1$ y $\rho_0 = q_1 = 2\rho_1/(1 + \rho_1^2)$, entonces $\alpha = \lim\limits_{k\to\infty} \alpha_k = 1 + \rho_1^2$, y para los k suficientemente grandes tenemos $\alpha_k \approx \alpha$. Por eso es natural estudiar *el método iterativo de tres capas estacionario*

$$By_{k+1} = \alpha (B - \tau A) y_k + (1 - \alpha) By_{k-1} + \alpha\tau f, \quad k = 1, 2, \ldots, \tag{3}$$

$$By_1 = (B - \tau A) y_0 + \tau f, \quad y_0 \in H$$

con parámetros constantes (estacionarios)

$$\alpha = 1 + \rho_1^2, \quad \tau = \tau_0 = \frac{2}{\gamma_1+\gamma_2}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}, \tag{4}$$

donde γ_1 y γ_2 son las constantes de equivalencia energética de los operadores autoconjugados D y $DB^{-1}A$:

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0, \quad DB^{-1}A = (DB^{-1}A)^*. \tag{5}$$

2. Estimación de la velocidad de convergencia. Para obtener una estimación de la velocidad de convergencia del método de tres capas estacionario pasemos de (3) al esquema de error equivalente $x_k = D^{1/2}z_k$:

$$x_{k+1} = \alpha (E - \tau C) x_k + (1 - \alpha) x_{k-1}, \quad k = 1, 2, \ldots$$
$$x_1 = (E - \tau C) x_0, \quad C = D^{1/2}B^{-1}AD^{-1/2}.$$

De aquí se deduce, que para cualquier $k \geqslant 0$, x_k se expresa mediante x_0 de la siguiente forma:

$$x_k = P_k (C) x_0, \tag{6}$$

donde el polinomio algebraico $P_k(t)$ correspondiente a $P_k(C)$ se determina por medio de las relaciones recurrentes

$$P_{k+1}(t) = \alpha(1 - \tau t) P_k(t) + (1 - \alpha) P_{k-1}(t), \quad k = 1, 2, \ldots, \tag{7}$$

$$P_1(t) = 1 - \tau t, \quad P_0(t) \equiv 1.$$

De (6) se deduce una estimación para la norma del error z_k en H_D:

$$\| z_k \|_D = \| x_k \| \leqslant \| P_k(C) \| \| x_0 \| = \| P_k(C) \| \| z_0 \|_D \tag{8}$$

Por eso es necesario valorar la norma del polinomio operacional $P_k(C)$ para el caso, cuando los parámetros α y τ son elegidos por las fórmulas (4). De las condiciones (5) se deduce que C es un operador autoconjugado en H y γ_1, γ_2 son sus cotas, por lo tanto,

$$\| P_k(C) \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} | P_k(t) |.$$

Expresemos el máximo del módulo del polinomio $P_k(t)$ sobre el segmento $[\gamma_1, \gamma_2]$. Para eso expresemos el polinomio $P_k(t)$ mediante los polinomios de Chebishev. A nosotros nos será más cómodo examinar $P_k(t)$ no sobre el segmento $[\gamma_1, \gamma_2]$, sino sobre el segmento estandar $[-1, 1]$. Poniendo

$$t = \frac{1 - \rho_0 x}{\tau_0}, \quad \tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

apliquemos el segmento $[\gamma_1, \gamma_2]$ sobre $[-1, 1]$. Entonces

$$P_k(t) = Q_k(x), \quad x \in [-1, 1],$$

$$\max_{\gamma_1 \leqslant t \leqslant \gamma_2} | P_k(t) | = \max_{|x| \leqslant 1} | Q_k(x) |.$$

Teniendo en cuenta la elección de los parámetros α y τ de acuerdo con (4), de (7) obtendremos las siguientes relaciones recurrentes para los polinomios $Q_k(x)$:

$$Q_{k+1}(x) = 2\rho_1 x Q_k(x) - \rho_1^2 Q_{k-1}(x), \quad k = 1, 2, \ldots,$$

$$Q_1(x) = \rho_0 x, \quad Q_0(x) \equiv 1.$$

De aquí con ayuda de la sustitución

$$Q_k = \rho_1^k R_k(x) \tag{9}$$

obtenemos fácilmente la relación recurrente estandar

$$R_{k+1}(x) = 2xR_k(x) - R_{k-1}(x), \quad k = 1, 2, \ldots,$$
$$R_1(x) = \rho_0 x/\rho_1, \quad R_0(x) \equiv 1. \tag{10}$$

A esta relación satisfacen el polinomio de Chebishev de primer género $T_k(x)$ con las condiciones iniciales $T_k(x) = x$, $T_0(x) \equiv 1$ y el polinomio de Chebishev de segundo género $U_k(x)$:

$$U_k(x) = \begin{cases} \dfrac{\operatorname{sen}((k+1)\arccos x)}{\operatorname{sen}(\arccos x)}, & |x| \leqslant 1, \\ \dfrac{\operatorname{sh}((k+1)\operatorname{Arch} x)}{\operatorname{sh}(\operatorname{Arch} x)}, & |x| \geqslant 1, \end{cases}$$

con las condiciones iniciales $U_1(x) = 2x$, $U_0(x) \equiv 1$. Empleando las propiedades indicadas de los polinomios $T_k(x)$ y $U_k(x)$ y la igualdad $\rho_0 = q_1 = 2\rho_1/(1+\rho_1^2)$, de (10) hallamos la expresión para el polinomio $R_k(x)$ mediante los polinomios de Chebishev

$$R_k(x) = \frac{2\rho_1^2}{1+\rho_1^2} T_k(x) + \frac{1-\rho_1^2}{1+\rho_1^2} U_k(x), \quad k \geqslant 0.$$

A continuación, utilizando las estimaciones conocidas

$$\max_{|x|\leqslant 1} |T_k(x)| = T_k(1) = 1,$$
$$\max_{|x|\leqslant 1} |U_k(x)| = U_k(1) = k+1,$$

obtenemos

$$\max_{|x|\leqslant 1} |R_k(x)| = R_k(1) = 1 + k(1-\rho_1^2)/(1+\rho_1^2).$$

De aquí, teniendo en cuenta los cambios hechos más arriba, hallamos la siguiente estimación para la norma del polinomio operacional $P_k(C)$:

$$\|P_k(C)\| \leqslant \rho_1^k (1 + k(1-\rho_1^2)/(1+\rho_1^2)). \tag{11}$$

Sustituyendo (11) en (8), obtenemos la estimación para la norma del error z_k en H_D:

$$|z_k|_D \leqslant \bar{q}_k |z_0|_D, \quad \bar{q}_k = \rho_1^k (1 + k(1-\rho_1^2)/(1+\rho_1^2)),$$

al mismo tiempo $q_k \to 0$ cuando $k \to \infty$ y $\bar{q}_{k+1} < \bar{q}_k$. Así está demostrado

TEOREMA 2. *El método iterativo de tres capas estacionario (3)-(5) converge en H_D, y para el error z_k es válida la esti-*

mación

$$\| z_k \|_D \leqslant \bar{q}_k \| z_0 \|_D, \quad \bar{q}_k = \rho_1^k (1 + k(1 - \rho_1^2)/(1 + \rho_1^2)).$$

OBSERVACION. Se puede mostrar, que $\lim\limits_{k\to\infty} q_k/\bar{q}_k = \lim\limits_{k\to\infty} \bar{q}_k/\rho_0^k = 0$, donde q_k está definido en el teorema 1. Por eso el método de tres capas estacionario converge con mayor rapidez que el método de iteración simple, pero más lentamente que el método de dos capas de Chebishev y el método semi-iterativo de Chebishev.

§ 4. Estabilidad de los métodos de dos y tres capas respecto a los datos a priori

1. **Planteamiento del problema.** Para la resolución aproximada de la ecuación operacional $Au = f$ en el capítulo VI fueron estudiados los métodos de dos capas de iteración simple y el método de Chebishev, y en los §§ 2 y 3 del capítulo VII fueron expuestos el método semi-iterativo de Chebishev y el método de tres capas estacionario.

Recordemos, que para calcular los parámetros iterativos en estos métodos se utiliza una información a priori determinada sobre los operadores del esquema iterativo. En el caso del operador autoconjugado $DB^{-1}A$ esta información tiene la forma de las constantes de equivalencia energética γ_1 y γ_2 de los operadores D y $DB^{-1}A$:

$$\gamma_1 (Dx, x) \leqslant (DB^{-1}Ax, x) \leqslant \gamma_2 (Dx, x), \quad \gamma_1 > 0. \quad (1)$$

En una serie de casos las constantes γ_1 y γ_2 pueden ser halladas con exactitud, es decir, existen tales elementos $x \in H$, para los cuales se alcanza la igualdad en (1). En otros casos para encontrar γ_1 y γ_2 se utilizan procesos auxiliares y estas constantes se hallan aproximadamente.

La utilización de la información inexacta a priori conduce a la disminución de la velocidad de convergencia, y en algunos casos a la divergencia del método. El objetivo del presente párrafo es estudiar la influencia de representar la información inexacta a priori sobre la velocidad de convergencia de los métodos iterativos enumerados más arriba.

Limitémonos al examen del caso autoconjugado, o sea, supongamos que el operador $DB^{-1}A$ es autoconjugado en H. Sean dados ciertos valores aproximados $\tilde{\gamma}_1$ y $\tilde{\gamma}_2$ en lugar

de los valores exactos γ_1 y γ_2 en las desigualdades (1). Examinemos los métodos de dos y tres capas, para los cuales eligiremos los parámetros iterativos de acuerdo con los $\widetilde{\gamma}_1$ y $\widetilde{\gamma}_2$ dados. Recordemos las fórmulas para los parámetros iterativos.

Para el esquema de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f, \quad k=0, 1, \ldots, \tag{2}$$

los parámetros del método de iteración se determinan por la fórmula

$$\tau_k=\widetilde{\tau}_0=2/(\widetilde{\gamma}_1+\widetilde{\gamma}_2), \quad k=1, 2, \ldots, \tag{3}$$

y los parámetros del método de Chebishev se construyen por la fórmula

$$\begin{aligned}&\tau_k=\widetilde{\tau}_0/(1+\widetilde{\rho}_0\mu_k), \quad \mu_k\in\mathfrak{M}_n^*, \quad k=1, 2, \ldots, n,\\ &\widetilde{\rho}_0=(1-\widetilde{\xi})/(1+\widetilde{\xi}), \quad \widetilde{\xi}=\widetilde{\gamma}_1/\widetilde{\gamma}_2.\end{aligned} \tag{4}$$

Para el esquema iterativo de tres capas

$$\begin{aligned}&By_{k+1}=\alpha_{k+1}(B-\tau_{k+1}A)y_k+(1-\alpha_{k+1})By_{k-1}+\alpha_{k+1}\tau_{k+1}f,\\ &k=1, 2, \ldots,\\ &By_1=(B-\tau_1A)y_0+\tau_1f\end{aligned} \tag{5}$$

los parámetros del método semi-iterativo de Chebishev se determinan por las fórmulas

$$\begin{aligned}&\tau_k\equiv\widetilde{\tau}_0, \quad \alpha_{k+1}=4/(4-\widetilde{\rho}_0^2\alpha_k),\\ &k=1, 2, \ldots, \alpha_1=2,\end{aligned} \tag{6}$$

mientras que los parámetros del método de tres capas estacionario se prefijan mediante las fórmulas

$$\begin{aligned}&\tau_k\equiv\widetilde{\tau}_0, \ \alpha_k\equiv 1+\widetilde{\rho}_1^2, \quad k=1, 2, \ldots\\ &\ldots, \widetilde{\rho}_1=\left(1-\sqrt{\widetilde{\xi}}\right)/\left(1+\sqrt{\widetilde{\xi}}\right).\end{aligned} \tag{7}$$

De la teoría general de los métodos iterativos, expuesta más arriba, se deduce que para el error $z_k=y_k-u$ de los métodos examinados son válidas las estimaciones:

1) para el método de iteración simple

$$\|z_n\|_D\leqslant\Big(\max_{\gamma_1\leqslant t\leqslant\gamma_2}|1+\widetilde{\tau}_0t|\Big)^n\|z_0\|_D; \tag{8}$$

2) para el método de dos capas de Chebishev y el método semi-iterativo de Chebishev

$$\| z_n \|_D \leqslant \widetilde{q}_n \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \left| T_n \left(\frac{1-\widetilde{\tau}_0 t}{\widetilde{\rho}_0} \right) \right| \| z_0 \|_D, \tag{9}$$

donde $\widetilde{q}_n = 2\widetilde{\rho}_1^n/(1+\widetilde{\rho}_1^{2n})$;

3) para el método de tres capas estacionario

$$\| z_n \|_D \leqslant \widetilde{\rho}_1^n \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \left| \frac{2\widetilde{\rho}_1^2}{1+\widetilde{\rho}_1^2} T_n \left(\frac{1-\widetilde{\tau}_0 t}{\widetilde{\rho}_0} \right) + \right.$$

$$\left. + \frac{1-\widetilde{\rho}_1^2}{1+\widetilde{\rho}_1^2} U_n \left(\frac{1-\widetilde{\tau}_0 t}{\widetilde{\rho}} \right) \right| \| z_0 \|_D. \tag{10}$$

Aquí $T_n(x)$ y $U_n(x)$ son los polinomios de Chebishev de primero y segundo género, γ_1, γ_2 son los valores exactos de las constantes de (1).

Las estimaciones citadas determinan la velocidad de convergencia de los métodos a examinar en el caso, cuando los parámetros iterativos se calculan por la información inexacta a priori.

2. Estimaciones de la velocidad de convergencia de los métodos. Estimemos ahora los máximos de los módulos de los polinomios, que forman parte de las estimaciones (8)-(10). Para eso haremos un cambio en (8)-(10), poniendo $x = (1 - \widetilde{\tau}_0 t)/\widetilde{\rho}_0$, y designaremos $a = (1 - \widetilde{\tau}_0 \gamma_2)/\widetilde{\rho}_0$, $b = (1 - \widetilde{\tau}_0 \gamma_1)/\widetilde{\rho}_0$. Entonces las estimaciones (8)-(10) tendrán la forma

$$\| z_n \|_D \leqslant \widetilde{\rho}_0^n \left(\max_{a \leqslant x \leqslant b} | x |^n \right) \| z_0 \|_D,$$

$$\| z_n \|_D \leqslant \widetilde{q}_n \max_{a \leqslant x \leqslant b} | T_n(x) | \| z_0 \|_D, \tag{11}$$

$$\| z_n \|_D \leqslant \widetilde{\rho}_1^n \max_{a \leqslant x \leqslant b} \left| \frac{2\widetilde{\rho}_1^2}{1+\widetilde{\rho}_1^2} T_n(x) + \frac{1-\widetilde{\rho}_1^2}{1+\widetilde{\rho}_1^2} U_n(x) \right| \| z_0 \|_D.$$

Examinemos primeramente el caso, cuando $\widetilde{\gamma}_1$ y $\widetilde{\gamma}_2$ son aproximaciones para γ_1 y γ_2 respectivamente, por debajo y por encima, es decir

$$\widetilde{\gamma}_1 \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \widetilde{\gamma}_2 \tag{12}$$

En este caso, como es fácil comprobar, se cumplirán las desigualdades $-1 \leqslant a \leqslant b \leqslant 1$. De (11) obtenemos, que la velocidad de convergencia del método de iteración simple se determinará por la magnitud $\widetilde{\rho}_0^n$, la del método de dos capas de Chebishev y del método semi-iterativo de Chebishev por la magnitud $\widetilde{q}_n$, y la del método de tres capas estacionario por la magnitud $\widetilde{\rho}_1^n (1 + n (1 - \widetilde{\rho}_1^2)/(1 + \widetilde{\rho}_1^2))$. Los métodos iterativos convergerán, pero la velocidad de convergencia disminuye.

Examinemos un ejemplo, para el cual se cumplen las condiciones (12). Dado que

$$\widetilde{\gamma}_1 = \gamma_1 (1 - \alpha), \quad \widetilde{\gamma}_2 = \gamma_2, \quad 0 \leqslant \alpha < 1.$$

En este caso $a = -1$ y $b < 1$. Por lo tanto para el error de los métodos a examinar de (11) obtenemos las siguientes estimaciones

$$\| z_n \|_D \leqslant \widetilde{\rho}_0^n \| z_0 \|_D,$$

$$\| z_n \|_D \leqslant \widetilde{q}_n \| z_0 \|_D,$$

$$\| z_n \|_D \leqslant \widetilde{\rho}_1^n (1 + n (1 - \widetilde{\rho}_1^2)/(1 + \widetilde{\rho}_1^2)) \| z_0 \|_D.$$

De las respectivas fórmulas para el número de iteraciones obtenemos, que para el método de iteración simple en el caso de la prefijación inexacta de γ_1 el número de iteraciones aumenta aproximadamente en $1/(1 - \alpha)$ veces en comparación con la prefijación exacta γ_1, en tanto que para el método de dos capas de Chebishev y el método semi-iterativo de Chebishev el número de iteraciones aumenta solamente en $1/\sqrt{1 - \alpha}$ veces.

Supongamos ahora que las condiciones (12) no son cumplidas. En este caso máx $(|a|, |b|) > 1$. Introduzcamos las siguientes notaciones

$$\frac{1}{\rho_0^*} = \text{máx}\,(|a|, |b|), \quad q_n^* = \frac{1}{T_n\left(\frac{1}{\rho_0^*}\right)} = \frac{2\rho_1^{*n}}{1 + \rho_1^{*2n}},$$

$$\rho_1^* = \frac{\rho_0^*}{1 + \sqrt{1 - \rho_0^{*2}}}.$$

Utilizando estas notaciones, y también la relación entre los polinomios de Chebishev de primero y segundo género

$$U_n (x) = (T_{n+1}^2 (x) - 1)^{1/2}/(T_1^2 (x) - 1)^{1/2}, \quad |x| \geqslant 1,$$

obtenemos

$$\max_{a\leqslant x\leqslant b} |x| = \frac{1}{\rho_0^*}, \quad \max_{a\leqslant x\leqslant b} |T_n(x)| = T_n\left(\frac{1}{\rho_0^*}\right) = \frac{1}{q_n^*} \leqslant \frac{1}{\rho_1^{*n}},$$

$$\max_{a\leqslant x\leqslant b} |U_n(x)| = U_n\left(\frac{1}{\rho_0^*}\right) = \frac{1-\rho_1^{*2(n+1)}}{\rho_1^{*n}(1-\rho_1^{*2})} \leqslant \frac{n+1}{\rho_1^{*n}}.$$

Sustituyendo estas estimaciones en (11), hallaremos

$$\| z_n \|_D \leqslant \left(\frac{\widetilde{\rho}_0}{\rho_0^*}\right)^n \| z_0 \|_D, \tag{13}$$

$$\| z_n \|_D \leqslant \frac{\widetilde{q}_n}{q_n^*} \| z_0 \|_D, \tag{14}$$

$$\| z_n \|_D \leqslant (\widetilde{\rho}_1/\rho_1^*)^n (1 + n(1-\widetilde{\rho}_1^2)/(1+\widetilde{\rho}_1^2)) \| z_0 \|_D. \tag{15}$$

Notemos, que si H es el espacio de dimensión finita, entonces se puede indicar una aproximación inicial y_0, para la cual se alcanzarán las igualdades en las estimaciones (13) y (14).

Hallemos ahora la condición, cumpliendo la cual se puede garantizar la convergencia de los métodos iterativos a examinar, construidos según una información inexacta a priori. Ya que la relación $\widetilde{q}_n/q_n^*$ tiende a cero cuando $n \to \infty$ solamente a la condición $\rho_1^* > \widetilde{\rho}_1$, mientras que esta condición es equivalente a la exigencia $\rho_0^* > \widetilde{\rho}_0$, entonces de (13)-(15) se deduce que los métodos iterativos convergerán, si se cumple la desigualdad

$$\widetilde{\rho}_0 < \rho_0^* \tag{16}$$

Utilizando las definiciones de ρ_0^*, a y b, obtenemos, que(16) tendrá lugar si $|1 - \widetilde{\tau}_0\gamma_1| < 1$, $|1 - \widetilde{\tau}_0\gamma_2| < 1$. Resolviendo estas desigualdades, hallamos

$$\widetilde{\gamma}_1 + \widetilde{\gamma}_2 > \gamma_2. \tag{17}$$

De esta forma, si se cumple la condición (17), entonces los métodos iterativos construidos por una información inexacta a priori, convergerán. De lo dicho más arriba se desprende que en el caso de un espacio de dimensión finita H la condición (17) es también necesaria para la convergencia de los métodos.

Definimos ahora el número real de iteraciones suficientes para alcanzar la exactitud prefijada ε. Como antes, designaremos mediante n el número de iteraciones para el caso cuando se da con exactitud la información a priori, mediante $\widetilde{n}$ designaremos el número teórico de iteraciones, calculado según las fórmulas de los correspondientes teoremas por una información inexacta a priori, y mediante n^* designemos el número real de iteraciones que es suficiente para

alcanzar la exactitud ε. De las fórmulas (13)-(15) se deduce que el número real de iteraciones n^* debe determinarse de las condiciones:

1) para el método de iteración simple de la condición $\widetilde{\rho}_0^n \leqslant \varepsilon \rho_0^{*n}$;

2) para el método de dos capas de Chebishev y el método semi-iterativo de Chebishev de la condición $\widetilde{q}_n \leqslant \varepsilon q_n^*$.

Es fácil cerciorarse de que tienen lugar las desigualdades $n^* \geqslant \widetilde{n}$, $n^* \geqslant n$ y al mismo tiempo el número de iteraciones $\widetilde{n}$ puede ser mayor o menor que n. Puesto que la única característica cuantitativa de un método iterativo, que puede ser calculada con anterioridad, es el número teórico de iteraciones n, entonces desde el punto de vista de la realización de los métodos es importante determinar o estimar en cuantas veces el número real de iteraciones n^* será mayor que $\widetilde{n}$. Para la comparación teórica de la calidad de los métodos iterativos es necesario estimar la relación n^*/n.

Obtendremos las estimaciones exigidas para un ejemplo. Sean $\widetilde{\gamma}_1$ y $\widetilde{\gamma}_2$ las aproximaciones para γ_1 y γ_2 respectivamente por encima y por debajo

$$\widetilde{\gamma}_1 = (1 + \alpha)\,\gamma_1, \quad \widetilde{\gamma}_2 = (1 - \alpha)\,\gamma_2, \quad \alpha \geqslant 0. \tag{18}$$

De la condición (17) y la exigencia natural $\widetilde{\gamma}_1 \leqslant \widetilde{\gamma}_2$ obtendremos, que los métodos convergerán, si se cumple la condición

$$\alpha < \text{mín}\,(\xi/(1 - \xi), \quad (1 - \xi)/(1 + \xi)), \quad \xi = \gamma_1/\gamma_2.$$

Para el ejemplo a examinar tendrán lugar las desigualdades $n^* \geqslant \geqslant n \geqslant \widetilde{n}$. En efecto, de (18) obtenemos

$$\widetilde{\xi} = \frac{\widetilde{\gamma}_1}{\widetilde{\gamma}_2} = \frac{1+\alpha}{1-\alpha}\,\xi \geqslant \xi$$

y, por consiguiente,

$$\widetilde{\rho}_0 \leqslant \rho_0 = (1-\xi)/(1+\xi), \quad \widetilde{\rho}_1 \leqslant \rho_1 = (1-\sqrt{\xi})/(1+\sqrt{\xi}), \quad \widetilde{q}_n \leqslant q_n;$$

De aquí se deduce que $n \geqslant \widetilde{n}$. Estimemos ahora las magnitudes que entran en las desigualdades (13)-(14). Ya que

$$\widetilde{\tau} = 2/(\widetilde{\gamma}_1 + \widetilde{\gamma}_2) = \tau_0/(1 - \alpha\rho_0) < \tau_0, \quad \tau_0 = 2/(\gamma_1 + \gamma_2),$$

entonces

$$1/\rho_0^* = \text{máx}\,(|\,a\,|, |\,b\,|) = |\,a\,| = (\widetilde{\tau}_0\gamma_2 - 1)/\widetilde{\rho}_0.$$

Omitiendo razonamientos no complejos, obtendremos

$$\widetilde{\rho}_0 = \frac{1-\widetilde{\xi}}{1+\widetilde{\xi}} = \frac{\rho_0-\alpha}{1-\alpha\rho_0},$$

$$\frac{\widetilde{\rho}_0}{\rho_0^*} = \widetilde{\tau}_0\gamma_2 - 1 = 1 - \frac{1-\frac{\alpha}{\xi}(1-\xi)}{1+\alpha}(1-\rho_0) =$$

$$= 1 - \frac{1-\frac{\alpha}{\xi}(1-\xi)}{1-\alpha\rho_0}(1-\rho_0),$$

$$\widetilde{\rho}_1 = \frac{1-\sqrt{\widetilde{\xi}}}{1+\sqrt{\widetilde{\xi}}} = \frac{\rho_0-\alpha}{1-\alpha\rho_0+\sqrt{(1-\alpha^2)(1+\rho_0^2)}},$$

$$\frac{\widetilde{\rho}_1}{\rho_1^*} = 1 - \frac{(1+\alpha)\sqrt{\xi}+\sqrt{1-\alpha^2}-\frac{\alpha}{\sqrt{\xi}}-\sqrt{\frac{\alpha}{\xi}[1-(1+\alpha)\xi]}}{(1+\alpha)\sqrt{\xi}+\sqrt{1-\alpha^2}} \times$$

$$\times (1-\widetilde{\rho}_1) = 1 -$$

$$- \frac{\left[(1+\alpha)\sqrt{\xi}+\sqrt{1-\alpha^2}-\frac{\alpha}{\sqrt{\xi}}-\sqrt{\frac{\alpha}{\xi}[1-(1+\alpha)\xi]}\right](1+\sqrt{\xi})}{1-\alpha+(1+\alpha)\xi+2\sqrt{(1-\alpha^2)\xi}} \times$$

$$\times (1-\rho_1).$$

Examinemos primeramente el método de iteración simple. Del teorema 2 del § 3 del cap. VI y de (13) obtenemos las estimaciones siguientes para el número de iteraciones n^*, $\widetilde{n}$ y n del método de iteración simple:

$$n = \frac{\ln\varepsilon}{\ln\rho_0} \approx \frac{\ln(1/\varepsilon)}{1-\rho_0}, \quad \widetilde{n} = \frac{\ln\varepsilon}{\ln\widetilde{\rho}_0} \approx \frac{\ln(1/\varepsilon)}{1-\widetilde{\rho}_0},$$

$$n^* = \frac{\ln\varepsilon}{\ln(\widetilde{\rho}_0/\rho_0^*)} = \frac{\ln(1/\varepsilon)}{1-\widetilde{\rho}_0/\rho_0^*}.$$

Sustituyendo aquí las expresiones obtenidas más arriba, hallaremos

$$\frac{n^*}{\widetilde{n}} \approx \frac{1+\alpha}{1-\frac{\alpha}{\xi}(1-\xi)}, \quad \frac{n^*}{n} \approx \frac{1-\alpha\rho_0}{1-\frac{\alpha}{\xi}(1-\xi)}.$$

Si $\alpha \approx c\xi$, donde $c < 1$, entonces de aquí obtendremos

$$n^* \approx \widetilde{n}/(1-c), \quad n^* \approx n/(1-c).$$

Así, si $\alpha \approx c\xi$, entonces el número real de iteraciones n^* para el método de iteración simple es $1/(1-c)$ veces mayor que el número teórico de iteraciones $\tilde{n}$, el cual se calcula según una información inexacta a priori.

Examinemos ahora el método de Chebishev y el método semi-iterativo de Chebishev. De la definición de $\tilde{q}_n$ y q_n^* obtenemos

$$\frac{\tilde{q}_n}{q_n^*} = \frac{\tilde{\rho}_1^n}{\rho_1^{*n}} \cdot \frac{1+\rho_1^{*2n}}{1+\tilde{\rho}_1^{2n}} \leqslant \frac{2\,(\tilde{\rho}_1/\rho_1^*)^n}{1+(\tilde{\rho}_1/\rho_1^*)^{2n}}.$$

Por esto para el número de iteraciones n^* hallaremos la siguiente estimación:

$$n^* = \frac{\ln(0{,}5\varepsilon)}{\ln(\tilde{\rho}_1/\rho_1^*)} \approx \frac{\ln(2/\varepsilon)}{1-\tilde{\rho}_1/\rho_1^*}.$$

Después, del teorema 1 del § 2 del cap. VI y del teorema 1 del § 2 del cap. VII obtendremos las estimaciones para n y $\tilde{n}$:

$$n = \frac{\ln(0{,}5\varepsilon)}{\ln\rho_1} \approx \frac{\ln(2/\varepsilon)}{1-\rho_1},$$

$$\tilde{n} = \frac{\ln(0{,}5\varepsilon)}{\ln\tilde{\rho}_1} \approx \frac{\ln(2/\varepsilon)}{1-\tilde{\rho}_1}.$$

Sustituyendo aquí las expresiones obtenidas más arriba para la relación $\tilde{\rho}_1/\rho_1^*$ y suponiendo, que $\alpha \approx c\xi$, hallaremos

$$n^*/\tilde{n} \approx 1/(1-\sqrt{c}), \quad n^*/n \approx 1/(1-\sqrt{c}).$$

De esta forma, si $\alpha \approx c\xi$, donde $c < 1$, entonces el número real de iteraciones n^* para el método de Chebishev y el método semi-iterativo de Chebishev es aproximadamente en $1/(1-\sqrt{c})$ veces mayor que el número teórico de iteraciones $\tilde{n}$ calculado según una información inexacta a priori.

Capítulo VIII

Métodos iterativos de tipo variacional

En este capítulo se examinan los métodos iterativos de dos y tres capas de tipo variacional. Para efectuar estos métodos no se exige ninguna información a priori sobre los operadores del esquema iterativo. En los §§ 1 y 2 se estudian los métodos gradientales de dos capas; en los §§ 3 y 4 los métodos de tres capas de direcciones conjugadas. El § 5 está dedicado a la aceleración de la convergencia de los métodos de dos capas en el caso autoconjugado.

§ 1. Métodos gradientales de dos capas

1. Planteamiento del problema sobre la elección de los parámetros iterativos. Para encontrar la solución aproximada de la ecuación operacional lineal

$$Au = f \tag{1}$$

con el operador no degenerado A, definido en un espacio de Hilbert real H, examinaremos el esquema iterativo implícito de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \tag{2}$$

con una aproximación inicial arbitraria $y_0 \in H$ y el operador no degenerado B.

El esquema iterativo (2) fue estudiado por nosotros anteriormente en el capítulo VI donde fueron construidos conjuntos de parámetros iterativos $\{\tau_k\}$ y fueron dadas estimaciones de la velocidad de convergencia de los correspondientes métodos iterativos (del método de Chebishev y del método de iteración simple).

Cualquier método iterativo de dos capas, construido a base del esquema (2), se caracteriza por los operadores A y B, por el espacio energético H_D, en el cual se demuestra la convergencia del método, y por el conjunto de parámetros iterativos τ_k. El problema fundamental de la teoría de los métodos iterativos es el problema sobre la elección óptima de los parámetros τ_k.

En el capítulo VI fueron construidos métodos iterativos, en los cuales los parámetros τ_k se eligieron de la condición de mínimo en H_D de la norma del operador de transición de una iteración a otra, o de la norma del operador de resolución. Un rasgo distintivo de los métodos iterativos construidos sobre este principio es la utilización de determinada información a priori acerca de los operadores del esquema iterativo para el cálculo de los parámetros τ_k.

El tipo de la información a priori exigida está determinada por las propiedades de los operadores A, B y D. Así, en el caso, cuando el operador $DB^{-1}A$ es autoconjugado en el espacio H, dicha información significa la prefijación de las constantes de equivalencia energética de los operadores D y $DB^{-1}A$, o sea, las constantes γ_1 y γ_2 de las desigualdades

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \ \gamma_1 > 0, \tag{3}$$

o las cotas del operador $DB^{-1}A$ en H_D.

Para el caso no autoconjugado se emplean o bien los dos números γ_1 y γ_2 de las desigualdades

$$\gamma_1 D \leqslant DB^{-1}A, \ (DB^{-1}Ax, B^{-1}Ax) \leqslant \gamma_2 (DB^{-1}Ax, x), \ \gamma_1 > 0, \tag{4}$$

o tres números —γ_1, γ_2 y γ_3, donde γ_1 y γ_2 son las constantes de las desigualdades (3), y γ_3 es o bien la constante de la desigualdad

$$\| 0{,}5 \ (DB^{-1}A - A^* (B^*)^{-1}D) \ x \|^2_{D^{-1}} \leqslant \gamma_3^2 (Dx, x) \tag{5}$$

o de la desigualdad

$$\| 0{,}5 \ (DB^{-1}A - A^* (B^*)^{-1}D) \ x \|^2_{D^{-1}} \leqslant \gamma_3 (DB^{-1}Ax, x). \tag{6}$$

En una serie de casos, encontrar las constantes γ_1, γ_2 y γ_3 con suficiente exactitud puede resultar un problema independiente complejo, que exige para su resolución la utilización de métodos computacionales especiales. Si la información a priori puede ser obtenida al precio de gastos

computacionales no grandes o si se exige resolver la serie de problemas (1) con miembros derechos diferentes, entonces es útil hallar de una vez la información a priori necesaria y después utilizar los métodos iterativos construidos en el capítulo VI. Este camino se puede recomendar, si el tiempo complementario, gastado para obtener la información a priori, es esencialmente menor que el tiempo de resolución de toda la serie de problemas (1).

En aquellos casos, cuando se exige resolver sólo un problema de (1). o cuando está prefijada una buena aproximación inicial, y el cálculo de las constantes γ_1, γ_2 y γ_3 es un proceso laborioso, se deben utilizar los métodos iterativos de tipo variacional, a cuyo examen nosotros pasamos.

En los métodos iterativos de dos capas de tipo variacional para calcular los parámetros τ_k no se exige ninguna información a priori sobre los operadores del esquema (2) (excepto de las condiciones de tipo general $A = A^* > 0$, $(DB^{-1}A)^* = DB^{-1}A$, etc.), y la construcción de estos métodos se basa en el siguiente principio. Si se da la aproximación y_k y y_{k+1} se encuentra por el esquema (2), entonces el parámetro iterativo τ_{k+1} se elige de la condición de mínimo en H_D de la norma del error $z_{k+1} = y_{k+1} - u$, donde u es la solución de la ecuación (1).

El nombre de métodos está relacionado con que la sucesión y_k, construida según la fórmula (2), en la cual los parámetros τ_k se eligen de la condición indicada más arriba, es una sucesión minimizante para el funcional cuadrático

$$I(y) = (D(y - u), y - u).$$

Este funcional en virtud de la definición positiva del operador D está acotado inferiormente y alcanza el mínimo, igual a cero, en la solución de la ecuación (1), es decir para $y = u$. La elección del parámetro τ_{k+1} de la condición indicada asegura la minimización local del funcional $I(y)$ al pasar de y_k a y_{k+1}, es decir, luego de un paso iterativo. En el caso de un esquema explícito ($B = E$) la transición de y_k a y_{k+1} se lleva a efecto según la fórmula

$$y_{k+1} = y_k - \tau_{k+1} r_k, \quad r_k = Ay_k - f.$$

Señalemos, que para un operador A autoconjugado y definido positivo la transición de y_k a y_{k+1} ocurre por la dirección de $-r_k$, la cual coincide con la dirección del antigradiente para el funcional $(A(y - u), y - u)$ en el

punto y_k. Es conocido, que por la dirección del antigradiente ocurre el mayor decrecimiento del valor del funcional. Por lo tanto estos métodos son llamados algunas veces métodos de descenso gradiental o sencillamente métodos gradientales. Conservaremos esta denominación también para los métodos de dos capas implícitas de tipo variacional.

Nuestro problema más cercano es encontrar el parámetro τ_{k+1} de la condición de mínimo de la norma del error $z_{k+1} = y_{k+1} - u$ en H_D.

2. **Fórmula para los parámetros iterativos.** Hallemos ahora una fórmula para calcular el parámetro iterativo τ_{k+1}, suponiendo que el operador A no es degenerado. Escribamos primeramente una ecuación para el error $z_k = y_k - u$, $k = 0, 1, \ldots$. Sustituyendo $y_k = z_k + u$ en el esquema (2), obtendremos

$$z_{k+1} = (E - \tau_{k+1}B^{-1}A)\, z_k, \quad k = 0, 1, \ldots, z_0 = y_0 - u.$$

El cambio $z_k = D^{-1/2}x_k$ permite pasar a una ecuación, que contiene solamente un operador

$$x_{k+1} = S_{k+1}x_k, \quad S_k = E - \tau_k C,$$
$$C = D^{-1/2}(DB^{-1}A)\, D^{-1/2}. \tag{7}$$

Utilizando la igualdad $\| z_k \|_D = \| x_k \|$, el problema planteado más arriba sobre la elección del parámetro τ_{k+1} se puede formular de la siguiente forma: elegir el parámetro τ_{k+1} de la condición de mínimo de la norma de x_{k+1} en el espacio H.

Resolvamos este problema. Calculemos la norma de x_{k+1}:

$$\| x_{k+1} \|^2 = ((E - \tau_{k+1}C)\, x_k, \quad (E - \tau_{k+1}C)\, x_k = \| x_k \|^2 -$$
$$- 2\tau_{k+1}(Cx_k, x_k) + \tau_{k+1}^2 (Cx_k, Cx_k) = (Cx_k, Cx_k) \times$$
$$\times \left[\tau_{k+1} - \frac{(Cx_k, x_k)}{(Cx_k, Cx_k)}\right]^2 + \| x_k \|^2 - \frac{(Cx_k, x_k)^2}{(Cx_k, Cx_k)}. \tag{8}$$

Como el operador A es no degenerado, entonces también lo es el operador C. Por eso para cualquier x_k tenemos $(Cx_k, Cx_k) > 0$, y el mínimo de la norma de x_{k+1} se alcanza para

$$\tau_{k+1} = \frac{(Cx_k, x_k)}{(Cx_k, Cx_k)}, \tag{9}$$

Sustituyendo (9) en (8), obtendremos

$$\| x_{k+1} \| = \rho_{k+1} \| x_k \|, \tag{10}$$

donde

$$\rho_{k+1}^2 = 1 - \frac{(Cx_k,\ x_k)^2}{(Cx_k,\ Cx_k)\,(x_k,\ x_k)}. \qquad (11)$$

De esta forma, la fórmula (9) determina el valor óptimo del parámetro iterativo τ_{k+1}. Sustituyendo $x_k = D^{1/2}z_k$ en (9), obtendremos

$$\tau_{k+1} = \frac{(DB^{-1}Az_k,\ z_k)}{(DB^{-1}Az_k,\ B^{-1}Az_k)}, \quad k = 0,\ 1,\ \ldots$$

Teniendo en cuenta, que $Az_k = Ay_k - Au = Ay_k - f = r_k$ es el error y $B^{-1}r_k = w_k$ la corrección, la fórmula para el parámetro τ_{k+1} se puede escribir en la forma siguiente

$$\tau_{k+1} = \frac{(Dw_k,\ z_k)}{(Dw_k,\ w_k)}, \qquad k = 0,\ 1,\ \ldots, \qquad (12)$$

y el esquema iterativo (2) se puede escribir en forma de una fórmula explícita para el cálculo de y_{k+1}:

$$y_{k+1} = y_k - \tau_{k+1}w_k, \quad k = 0,\ 1,\ \ldots \qquad (13)$$

El algoritmo que realiza el método construido se puede describir de la forma siguiente:

1) por y_k prefijado se calcula el error $r_k = Ay_k - f$,
2) se resuelve la ecuación para la corrección $Bw_k = r_k$,
3) por la fórmula (12) se calcula el parámetro τ_{k+1},
4) por la fórmula (13) se encuentra la nueva aproximación y_{k+1}.

Las fórmulas (12) aún no están aptas para los cálculos, ya que en el proceso de las iteraciones ellas contienen al mismo tiempo las magnitudes conocidas r_k y w_k y el error desconocido z_k. En el § 2, eligiendo un operador concreto D, nosotros obtendremos fórmulas para los parámetros τ_k, las cuales contendrán sólo magnitudes conocidas. Y ahora nosotros pasamos a la obtención de una estimación de la velocidad de convergencia del método iterativo construido.

3. Estimación de la velocidad de convergencia. Estimemos ahora la velocidad de convergencia de los métodos gradientales de dos capas. Puesto que el parámetro iterativo τ_{k+1} se elige de la condición de mínimo de la norma del error z_{k+1} en H_D, lo cual es equivalente a la condición

de mínimo de la norma de x_{k+1} en H, entonces de (7) obtendremos

$$\| x_{k+1} \| = \min_{\tau_{k+1}} \| S_{k+1} x_k \| \leqslant \min_{\tau_{k+1}} \| S_{k+1} \| \| x_k \| =$$

$$= \min_{\tau} \| E - \tau C \| \| x_k \| = \rho \| x_k \|, \quad \rho = \min_{\tau} \| E - \tau C \|.$$

Comparando esta estimación con la igualdad (10), encontramos

$$\rho_k \leqslant \rho \leqslant 1, \quad k = 1, 2, \ldots . \tag{14}$$

De (10) y (14) se deduce la estimación $\| x_{k+1} \| \leqslant \leqslant \rho \| x_k \|$, y en virtud del cambio hecho $x_k = D^{1/2} z_k$, de aquí se infiere una estimación para la norma del error z_n en el espacio energético H_D:

$$\| z_n \|_D \leqslant \rho^n \| z_0 \|_D, \quad \rho = \min_{\tau} \| E - \tau C \|. \tag{15}$$

Si se cumple la condición $\rho < 1$, entonces el método gradiental de dos capas converge en H_D. De la estimación (15) se deduce que para disminuir en $1/\varepsilon$ veces la norma del error inicial en H_D es suficiente cumplimentar $n \geqslant n_0(\varepsilon)$ iteraciones, donde

$$n_0(\varepsilon) = \ln \varepsilon / \ln \rho. \tag{16}$$

Así, la velocidad de convergencia de un método gradiental de dos capas está determinada por la magnitud ρ. Recordemos, que en el capítulo VI estudiando del método de iteración simple bajo diferentes suposiciones respecto al operador C fueron obtenidas estimaciones para ρ. La magnitud ρ determina la velocidad de convergencia del método de iteración simple. Por esto de la estimación (15) obtenida aquí se deduce, que todo método gradiental de dos capas converge no más lentamente que el correspondiente método de iteración simple.

Citemos las estimaciones para ρ, obtenidas en los §§ 3 y 4 del capítulo VI siendo diferentes suposiciones respecto a los operadores A, B y D.

1. Si el operador $DB^{-1}A$ es autoconjugado en H, y γ_1 y γ_2 son las constantes de las desigualdades (3), entonces

$$\rho = (1 - \xi)/(1 + \xi), \quad \xi = \gamma_1/\gamma_2. \tag{17}$$

2. Sea el operador $DB^{-1}A$ autoconjugado en H:

a) si son cumplidas las condiciones (4), entonces

$$\rho = \sqrt{1-\xi}, \quad \xi = \gamma_1/\gamma_2; \tag{18}$$

b) si se cumplen las condiciones (3), (5), entonces

$$\rho = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{1-\varkappa}{1+\varkappa}\frac{\gamma_1}{\gamma_2}, \quad \varkappa = \frac{\gamma_3}{\sqrt{\gamma_1\gamma_2+\gamma_3^2}}. \tag{19}$$

Así, está demostrado

TEOREMA 1. *Si para el esquema (2) el método de iteración simple converge, entonces converge el método gradiental de dos capas (2), (12). Con esto para el error z_n es válida la estimación*

$$\| z_n \|_D \leqslant \rho^n \| z_0 \|_D,$$

donde ρ está definido en (17), si el operador $DB^{-1}A$ es autoconjugado en H y son cumplidas las condiciones (3), ρ está definido en (18), si para el operador no autoconjugado $DB^{-1}A$ son cumplidas las condiciones (4) y está definido en (19), si se cumplen las condiciones (3), (15). La estimación para el número de iteraciones está dada en (16).

OBSERVACION. Si la ecuación (1) se examina en un espacio de Hilbert complejo, entonces el parámetro iterativo τ_{k+1} debe ser elegido por la fórmula

$$\tau_{k+1} = \frac{\operatorname{Re}(Dw_k, z_k)}{(Dw_k, w_k)}, \quad k = 0, 1, \ldots .$$

El teorema 1 conserva su fuerza, solamente las condiciones (3) y (4) deben ser cambiadas por las desigualdades

$$\gamma_1 (Dx, x) \leqslant \operatorname{Re}(DB^{-1}Ax, x) \leqslant \gamma_2 (Dx, x),$$
$$\gamma_1 (Dx, x) \leqslant \operatorname{Re}(DB^{-1}Ax, x),$$
$$(DB^{-1}Ax, B^{-1}Ax) \leqslant \gamma_2 \operatorname{Re}(DB^{-1}Ax, x),$$

donde Re z es la parte real del número complejo z.

4. **Inmejorabilidad de la estimación en el caso autoconjugado.** Mostremos que en la clase de las aproximaciones iniciales arbitrarias y_0 en el caso del operador autoconjugado $DB^{-1}A$ en el espacio de dimensión finita H la estimación a priori del error del método iterativo (2), (12), obtenida en el teorema 1, es inmejorable. Para eso es suficiente indicar una aproximación inicial x_0, bajo la cual para la solución de la ecuación (7) tenga lugar la igualdad $\| x_{k+1} \| = \rho \| x_k \|$, donde ρ está definido en (17).

Hallemos la aproximación inicial buscada x_0. Sea H un espacio de dimensión finita ($H = H_N$). Ya que el operador $DB^{-1}A$ es autoconjugado en H, entonces el operador $C = D^{-1/2}(DB^{-1}A)D^{-1/2}$ también es autoconjugado en H. Por consiguiente, existe un sistema completo de funciones propias $v_1, v_2, \ldots, v_N$ del operador C. Designemos mediante λ_k el valor propio del operador C, correspondiente a la función propia v_k, así que $Cv_k = \lambda_k v_k$, $k = 1, 2, \ldots, N$. Sean $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_N$. Ya que las desigualdades (3) son equivalentes a las desigualdades

$$\gamma_1 E \leqslant C \leqslant \gamma_2 E, \quad \gamma_1 > 0,$$

entonces en (3) en calidad de γ_1 y γ_2 se pueden tomar λ_1 y λ_N. Con esto ρ, definido en (17), se puede escribir en la siguiente forma: $\rho = (\lambda_N - \lambda_1)/(\lambda_N + \lambda_1)$. Elijamos la aproximación inicial

$$x_0 = \sqrt{\lambda_N} v_1 + \sqrt{\lambda_1} v_N. \tag{20}$$

Entonces $Cx_0 = \lambda_1 \sqrt{\lambda_N} v_1 + \lambda_N \sqrt{\lambda_1} v_N$. Utilizando la ortonormalidad del sistema de funciones propias $v_1, v_2, \ldots, v_N$, obtenemos

$$\begin{aligned} (x_0, x_0) &= \lambda_1 + \lambda_N, \\ (Cx_0, x_0) &= 2\lambda_1\lambda_N, \\ (Cx_0, Cx_0) &= \lambda_1\lambda_N(\lambda_1 + \lambda_N). \end{aligned}$$

Sustituyendo estos valores en (9), (11), obtenemos

$$\tau_1 = 2/(\lambda_1 + \lambda_N), \quad \rho_1 = (\lambda_N - \lambda_1)/(\lambda_N + \lambda_1) = \rho.$$

De (10) se desprende la igualdad $\| x_1 \| = \rho \| x_0 \|$, y de (7) hallaremos x_1:

$$x_1 = \rho(\sqrt{\lambda_N} v_1 - \sqrt{\lambda_1} v_N).$$

Cálculos posteriores dan

$$\begin{aligned} Cx_1 &= \rho(\lambda_1 \sqrt{\lambda_N} v_1 - \lambda_N \sqrt{\lambda_1} v_N), \\ (x_1, x_1) &= \rho^2 (x_0, x_0), \\ (Cx_1, x_1) &= \rho^2 (Cx_0, x_0), \\ (Cx_1, Cx_1) &= \rho^2 (Cx_0, Cx_0). \end{aligned}$$

Por eso

$$\tau_2 = \frac{(Cx_1,\ x_1)}{(Cx_1,\ Cx_1)} = \frac{(Cx_0,\ x_0)}{(Cx_0,\ Cx_0)} = \tau_1;$$

$$\rho_2^2 = 1 - \frac{(Cx_1,\ x_1)^2}{(Cx_1,\ Cx_1)\,(x_1,\ x_1)} = 1 - \frac{(Cx_0,\ x_0)^2}{(Cx_0,\ Cx_0)\,(x_0,\ x_0)} =$$

$$= \rho_1^2 = \rho^2.$$

Por consiguiente, $\| x_2 \| = \rho \| x_1 \|$. Además, $x_2 = x_1 - \tau_2 C x_1 = \rho^2 x_0$, es decir x_2 es proporcional a x_0. De aquí se deduce inmediatamente, que $\tau_3 = \tau_2 = \tau_1$, $\rho_3 = \rho$ y $x_3 = \rho^2 x_1$. Por eso para todo k:

$$\tau_k \equiv 2/(\lambda_1 + \lambda_N), \quad \rho_k \equiv \rho = (\lambda_N - \lambda_1)/(\lambda_N + \lambda_1),$$

$$\| x_{k+1} \| = \rho \| x_k \|.$$

La afirmación está demostrada.

Así, hemos mostrado, que si la aproximación inicial es elegida por la fórmula (20), entonces en el método gradiental de dos capas todos los parámetros τ_k son iguales y coinciden con el parámetro del método de iteración simple (véase el § 3 del cap. VI), los errores al cabo de una iteración son proporcionales, y la velocidad de convergencia es la más lenta.

Señalemos que esta convergencia lenta del método tiene lugar únicamente para una aproximación inicial «mala» especial. En el caso de una «buena» aproximación inicial la velocidad de convergencia del método puede ser considerablemente mayor. Un estudio más detallado del carácter de la variación de la velocidad de convergencia del método será mostrado en el punto siguiente, y aquí nosotros examinaremos un ejemplo, que ilustra la observación citada más arriba.

Mostraremos que si en calidad de aproximación inicial x_0 tomamos toda función propia v_m, entonces el método gradiental de dos capas convergerá durante una iteración.

En efecto, sea $x_0 = v_m$. Entonces cálculos no complejos dan

$$Cx_0 = \lambda_m v_m = \lambda_m x_0,\ (Cx_0,\ x_0) = \lambda_m\,(x_0,\ x_0),$$

$$(Cx_0,\ Cx_0) = \lambda_m^2\,(x_0,\ x_0),\ \tau_1 = 1/\lambda_m,\ \rho_1 = 0,$$

es decir, $x = 0$ o $y_1 = u$.

Esta cualitativamente nueva propiedad de los métodos gradientales de dos capas —la posibilidad de aumentar la

velocidad de convergencia en el caso, cuando está dada una «buena» aproximación inicial—, diferencia dichos métodos de los métodos iterativos de dos capas examinados en el capítulo VI, rigurosamente orientados a la más mala aproximación inicial.

5. Propiedad asintótica de los métodos gradientales en el caso autoconjugado. Examinemos ahora una propiedad asintótica de los métodos gradientales de dos capas, que ellos poseen en el caso cuando el operador $DB^{-1}A$ es autoconjugado. Esta propiedad consiste en que la sucesión $\{\rho_k\}$, definida en (11), es creciente. Puesto que la magnitud ρ_k define la velocidad de decrecimiento de la norma del error al pasar de la k-ésima iteración a la $(k+1)$-ésima, entonces la presencia de la propiedad indicada conduce a la disminución de la velocidad de decrecimiento de la norma del error z_n para los n grandes en comparación con el inicio del proceso de iteración. Al mismo tiempo para los n suficientemente grandes la velocidad de convergencia de los métodos gradientales se hace prácticamente igual que para el método de iteración simple.

Será mostrado, que para grandes números de iteraciones los errores, examinados después de una iteración, se hacen casi proporcionales. Utilizando este hecho, nosotros construiremos un método aproximado para encontrar las constantes γ_1 y γ_2 de los desigualdades (3) y en el § 5 construiremos un proceso de aceleración de la convergencia de los métodos gradientales de dos capas.

Así pues, sean autoconjugados en H el operador $DB^{-1}A$, y junto con él el operador C. Mostraremos que la sucesión $\{\rho_k\}$ es creciente. De (10) se deducen las igualdades

$$\| x_{k+2} \| = \rho_{k+2} \| x_{k+1} \|, \quad x_{k+1} \| = \rho_{k+1} \| x_k \|.$$

Calculemos la norma de la diferencia $x_{k+2} - \rho_{k+2}\rho_{k+1}x_k$:

$$\| x_{k+2} - \rho_{k+2}\rho_{k+1}x_k \|^2 =$$
$$= \| x_{k+2} \|^2 - 2\rho_{k+2}\rho_{k+1}(x_{k+2}, x_k) + \rho_{k+2}^2\rho_{k+1}^2 \| x_k \|^2 =$$
$$= 2(\| x_{k+2} \|^2 - \rho_{k+2}\rho_{k+1}(x_{k+2}, x_k)). \qquad (21)$$

Calculemos por separado el producto escalar (x_{k+2}, x_k). De la ecuación (7) hallaremos

$$x_{k+2} = x_{k+1} - \tau_{k+2}Cx_{k+1}, \quad x_k = x_{k+1} + \tau_{k+1}Cx_k. \qquad (22)$$

Multiplicando la última igualdad escalarmente por Cx_k y teniendo en cuenta (9), obtendremos

$$(Cx_k,\ x_k) = (x_{k+1},\ Cx_k) + \tau_{k+1}\,(Cx_k,\ Cx_k) = \\ = (x_{k+1},\ Cx_k) + (Cx_k,\ x_k).$$

Por consiguiente, para cualquier k tiene lugar la igualdad

$$(x_{k+1},\ Cx_k) = 0, \tag{23}$$

y en virtud de la autoconjugación de C se tiene la igualdad $(Cx_{k+1},\ x_k) = 0$.

De (22) y (23) obtenemos

$$(x_{k+2},\ x_k) = (x_{k+1} - \tau_{k+2}Cx_{k+1},\ x_k) = (x_{k+1},\ x_k) = \\ = (x_{k+1},\ x_{k+1} + \tau_{k+1}Cx_k) = \| x_{k+1} \|^2.$$

Sustituyendo la igualdad obtenida en (21), hallaremos

$$\| x_{k+2} - \rho_{k+2}\rho_{k+1}x_k \|^2 = 2\left(1 - \frac{\rho_{k+1}}{\rho_{k+2}}\right) \| x_{k+2} \|^2. \tag{24}$$

De (24) se sigue que o bien $\rho_{k+2} > \rho_{k+1}$ ó $\rho_{k+1} = \rho_{k+2} = \bar{\rho}$ y $x_{k+2} = \bar{\rho}^2 x_k$. En el último caso, es evidente que para todos los $n \geqslant k$ se cumplirán las igualdades

$$\rho_{n+1} = \bar{\rho}, \qquad x_{n+2} = \bar{\rho}^2 x_n, \tag{25}$$

es decir, la sucesión ρ_k sale a un valor límite.

De esta forma está mostrado que la sucesión $\{\rho_k\}$ realmente es creciente. En el punto 3 de este párrafo fue mostrado que esta sucesión está acotada por encima y, por lo tanto, posee un límite. Por esto para los números k bastante grandes tendrá lugar la igualdad aproximada $\rho_{k+1} \approx \rho_{k+2}$ y, por consiguiente, $x_{k+2} \approx \rho_{k+2}\rho_{k+1}x_k$, es decir, los errores después de una iteración serán casi proporcionales.

Examinemos que se deduce de la salida de la sucesión ρ_k a un valor límite. En este caso se cumplen las igualdades (25), o sea, $x_{n+2} = \bar{\rho}^2 x_n$. Sea el espacio H de dimensión finita, y $v_1, v_2, \ldots, v_N$ el sistema de funciones propias del operador C. Desarrollemos x_n según las funciones propias

$$x_n = \sum_{k=1}^{N} \alpha_k^{(n)} v_k. \tag{26}$$

De la ecuación (7) obtenemos

$$x_{n+2} = (E - \tau_{n+2}C)(E - \tau_{n+1}C)\, x_n = \sum_{k=1}^{N} (1 - \tau_{n+2}\lambda_k)(1 - \tau_{n+1}\lambda_k)\, \alpha_k^n v_k.$$

Ya que $x_{n+2} = \overline{\rho}^2 x_n$, entonces esto significa que para todos los números k para los cuales $\alpha_k^{(n)} \neq 0$, debe cumplirse la igualdad

$$(1 - \tau_{n+2}\lambda_k)(1 - \tau_{n+1}\lambda_k) = \overline{\rho}^2.$$

De aquí se sigue que en el desarrollo (26) están presentes las funciones propias, correspondientes sólo a dos valores propios diferentes (cada uno de los cuales puede ser múltiple). Supongamos que ellos sean λ_i y λ_j. Entonces λ_i y λ_j son las raíces de la ecuación

$$(1 - \tau_{n+2}\lambda)(1 - \tau_{n+1}\lambda) = \overline{\rho}^2. \tag{27}$$

Conociendo τ_{n+1}, τ_{n+2} y $\overline{\rho}$ de esta ecuación se pueden hallar los valores propios λ_i y λ_j.

No deteniéndonos en detalles, señalemos que si en el desarrollo del error inicial x_0 están presentes funciones propias correspondientes a los valores propios mínimo λ_1 y máximo λ_N del operador C, entonces en el caso de salida de la sucesión ρ_n a un valor límite en el desarrollo (26) quedarían solamente estas funciones propias. Por eso, resolviendo la ecuación (27), hallaremos λ_1 y λ_N.

La salida de la sucesión $\{\rho_n\}$ a un valor límite para n finito es un caso excepcional. En el caso general se puede únicamente afirmar que para n suficientemente grande

$$\rho_{n+1} \approx \rho_{n+2} \text{ y } x_{n+2} \approx \rho_{n+2}\rho_{n+1}x_n.$$

La presencia de una igualdad aproximada permite presuponer que para n suficientemente grande las raíces de la ecuación

$$(1 - \tau_{n+2}\lambda)(1 - \tau_{n+1}\lambda) = \rho_{n+2}\rho_{n+1} \tag{28}$$

serán buenas aproximaciones para λ_1 y λ_N y, por lo tanto, también para γ_1 y γ_2 de las desigualdades (3).

Describamos este método para encontrar valores aproximados de γ_1 y γ_2. Por el esquema iterativo (2) con $f = 0$ se realizan $n + 2$ iteraciones con los parámetros τ_{k+1}, definidos en (12). Como para $f = 0$ la solución de la ecuación

(1) es el cero ($u = 0$), entonces $z_k = y_k$ y, por consiguiente, ρ_{k+1} se puede hallar por la fórmula

$$\rho_{k+1} = \frac{\| z_{k+1} \|_D}{\| z_k \|_D} = \frac{\| y_{k+1} \|_D}{\| y_k \|_D}.$$

Calculando τ_{n+1}, τ_{n+2}, ρ_{n+1} y ρ_{n+2} se resuelve la ecuación (28). Las raíces de esta ecuación son aproximaciones para γ_1 por encima y para γ_2 por debajo.

En el § 5 será expuesto un ejemplo que ilustra el método propuesto para encontrar γ_1 y γ_2.

§ 2. Ejemplos de métodos gradientales de dos capas

1. Método del descenso más rápido. En el § 1 fueron estudiadas las propiedades generales de los métodos iterativos de dos capas de tipo variacional empleados para encontrar la solución aproximada de la ecuación operacional lineal

$$Au = f \tag{1}$$

con el operador no degenerado A. Las aproximaciones iterativas se calculan por el esquema de dos capas

$$B \frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, x_0 \in H, \tag{2}$$

y los parámetros iterativos τ_k se encuentran por la fórmula

$$\tau_{k+1} = \frac{(Dw_k, z_k)}{(Dw_k, w_k)}, \quad k = 0, 1, \ldots, \tag{3}$$

donde $w_k = B^{-1}r_k$ es la corrección, $r_k = Ay_k - f$ el defecto, y $z_k = y_k - u$ el error de las iteraciones. La elección del parámetro τ_{k+1} por la fórmula (3) asegura el mínimo de la norma del error z_{k+1} en H_D al pasar de y_k a y_{k+1}.

Examinemos ahora los casos particulares de los métodos gradientales de dos capas. Cada método concreto se determina por la elección del operador D y posee su dominio de aplicabilidad. El operador D se eligirá de manera tal, que en la fórmula (3) para el parámetro iterativo τ_{k+1} entren solamente cantidades conocidas en el proceso de iteraciones.

Comenzaremos el examen de los ejemplos por el método de descenso más rápido. Este método se puede aplicar únicamente en el caso de un operador A autoconjugado y definido positivo.

Sea el operador A autoconjugado y definido positivo en H. *El método del descenso más rápido* se caracteriza por la siguiente elección del operador D: $D = A$. El operador B debe ser definido positivo en H. Teniendo en cuenta la relación $Az_k = Ay_k - f = r_k$ y $A = A^*$, de (3) obtendremos una fórmula para el parámetro iterativo τ_{k+1} en el método implícito del descenso más rápido

$$\tau_{k+1} = \frac{(r_k, w_k)}{(Aw_k, w_k)}, \quad k = 0, 1, \ldots$$

Para el caso del esquema de dos capas explícito (2) ($B = E$) obtenemos $w_k = B^{-1}r_k = r_k$, y la fórmula para τ_{k+1} toma la forma

$$\tau_{k+1} = \frac{(r_k, r_k)}{(Ar_k, r_k)}, \quad k = 0, 1, \ldots .$$

En el método del descenso más rápido se minimiza la norma del error z_{k+1} en el espacio energético H_A: $\| z_k \|_A = (Az_k, z_k)^{1/2}$. Las condiciones de convergencia del método están formuladas en el teorema 1, del cual se deducen las estimaciones

$$\| z_n \|_A \leqslant \rho^n \| z_0 \|_A, \quad n \geqslant n_0(\varepsilon) = \ln \varepsilon / \ln \rho.$$

El valor de la magnitud ρ se determina por las propiedades de los operadores A y B y por el volumen de información a priori respecto a ellos. Notemos que la exigencia de autoconjugación del operador $DB^{-1}A = AB^{-1}A$ para el método dado es equivalente a la exigencia de autoconjugación del operador B. Por eso

1) si $B = B^*$ y se cumplen las condiciones (3) del § 1 ó las condiciones equivalentes (véase el cap. VI, § 2, punto 3)

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0,$$

entonces

$$\rho = (1 - \xi)/(1 + \xi), \quad \xi = \gamma_1/\gamma_2;$$

2) si $B \neq B^*$ y se cumplen las condiciones (4) del § 1 ó las condiciones equivalentes (véase el cap. VI, § 4, punto 2)

$$\gamma_1 (Bx, A^{-1}Bx) \leqslant (Bx, x), \quad (Ax, x) \leqslant \gamma_2 (Bx, x),$$
$$\gamma_1 > 0,$$

entonces

$$\rho = \sqrt{1 - \xi}, \quad \xi = \gamma_1/\gamma_2.$$

Señalemos que si $B = B^*$, entonces el método del descenso más rápido posee la propiedad asintótica.

2. **Método de los defectos mínimos.** Este método se puede aplicar para todo operador A no autoconjugado y no degenerado. La definición positiva de cada uno de los operadores A y B por separado no se supone, se exige solamente que el operador B^*A sea definido positivo. *El método de los defectos mínimos* se determina por la siguiente elección del operador D: $D = A^*A$.

La fórmula (3) para el parámetro iterativo τ_{k+1} en el método de los defectos mínimos posee la forma

$$\tau_{k+1} = \frac{(Aw_k,\ r_k)}{(Aw_k,\ Aw_k)}, \quad k = 0,\ 1,\ \ldots$$

En el caso del esquema explícito (2) ($B = E$) necesita la definición positiva del operador A, y la fórmula para τ_{k+1} tiene la forma

$$\tau_{k+1} = \frac{(Ar_k,\ r_k)}{(Ar_k,\ Ar_k)}, \quad k = 0,\ 1,\ \ldots$$

El nombre del método está relacionado con que en él se minimiza la norma del defecto (error de las imágenes iteradas). En efecto, para el operador D indicado tenemos

$$\| z_k \|_D^2 = (Dz_k,\ z_k) = (A^*Az_k,\ z_k) = \| Az_k \|^2 = \| r_k \|^2.$$

Por lo tanto, para el método examinado la norma del error en H_D es igual a la norma del defecto, el cual se puede calcular en el proceso de las iteraciones y emplear para controlar la terminación de las mismas.

Del teorema 1 se deducen las estimaciones de convergencia del método

$$\| r_n \| \leqslant \rho^n \| r_0 \|, \quad n \geqslant n_0(\varepsilon) = \ln \varepsilon / \ln \rho.$$

El operador $DB^{-1}A = A^*AB^{-1}A$ será autoconjugado en H, si lo es el operador AB^{-1}, lo que es equivalente a la exigencia de autoconjugación del operador B^*A. Si esta exigencia está cumplida, entonces de las condiciones (3) del § 1, las cuales en este caso tienen la forma

$$\gamma_1 (Ay,\ Ay) \leqslant (AB^{-1}Ay,\ Ay) \leqslant \gamma_2 (Ay,\ Ay),\ \gamma_1 > 0,$$

o después del cambio $y = A^{-1}Bx$

$$\gamma_1 (Bx,\ Bx) \leqslant (Ax,\ Bx) \leqslant \gamma_2 (Bx,\ Bx),\ \gamma_1 > 0, \quad (4)$$

y del teorema 1 se deduce que

$$\rho = (1 - \xi)/(1 + \xi), \quad \xi = \gamma_1/\gamma_2.$$

Señalemos que la condición $\gamma_1 > 0$ será cumplida, si se cumple la condición exigida más arriba de definición positiva del operador B^*A. Las condiciones de autoconjugación y definición positiva del operdor B^*A se cumplirán, por ejemplo, bajo las siguientes suposiciones:

$$A = A^* > 0, \quad B = B^* > 0, \quad AB = BA.$$

En este caso las desigualdades (4) son equivalentes a otras más simples. En realidad, poniendo $x = B^{-1/2}y$ en (4) y utilizando la conmutabilidad de los operadores A y B, obtenemos

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0. \tag{5}$$

Las condiciones de autoconjugación y definición positiva del operador B^*A se cumplirán automáticamente en el caso, cuando el operador B tiene la forma $B = (A^*)^{-1}B_0$, donde B_0 es un operador autoconjugado y definido positivo. En este caso en lugar de las desigualdades (5) es necesario utilizar las desigualdades

$$\gamma_1 B_0 \leqslant A^*A \leqslant \gamma_2 B_0, \quad \gamma_1 > 0, \tag{6}$$

y en la fórmula para el parámetro τ_{k+1} la corrección w_k se puede encontrar de la ecuación $B_0 w_k = A^* r_k$.

Si el operador B^*A no es autoconjugado en H, entonces de las condiciones (4) del § 1 o de las condiciones equivalentes

$$\gamma_1 (Bx, Bx) \leqslant (Ax, Bx), \quad (Ax, Ax) \leqslant \gamma_2 (Ax, Bx), \quad \gamma_1 > 0$$

y del teorema 1 se sigue que $\rho = \sqrt{1 - \xi}$, $\xi = \gamma_1/\gamma_2$.

3. Método de las correcciones mínimas. Este método se puede aplicar para resolver la ecuación (1) con un operador A no autoconjugado pero definido positivo. Se exige que el operador B sea autoconjugado, definido positivo y acotado. *El método de las correcciones* mínimas se determina por la siguiente elección del operador D: $D = A^*B^{-1}A$.

La fórmula (3) para el parámetro iterativo τ_{k+1} en el método de las correcciones mínimas tiene la forma

$$\tau_{k+1} = \frac{(Aw_k, w_k)}{(B^{-1}Aw_k, Aw_k)}, \quad k = 0, 1, \ldots$$

En el caso del esquema explícito (2) ($B = E$) los métodos de correcciones mínimas y de defectos mínimos coinciden.

En el método de las correcciones mínimas se minimiza la norma del error en H_B. En efecto, para el operador D elegido obtenemos

$$\| z_k \|_D^2 = (Dz_k, z_k) = (A^*B^{-1}Az_k, z_k) = (w_k, r_k) = \\ = (Bw_k, w_k) = \| w_k \|_B^2.$$

La norma de la corrección en H_B puede ser calculada en el proceso de las iteraciones y utilizada para controlar la terminación de las mismas.

Del teorema 1 se deducen los estimados de la convergencia del método

$$\| w_n \|_B \leqslant \rho^n \| w_0 \|_B, \quad n \geqslant n_0(\varepsilon) = \ln \varepsilon / \ln \rho.$$

El operador $DB^{-1}A = A^*B^{-1}AB^{-1}A$ es autoconjugado en H simultáneamente con el operador A. Por eso:

1) si $A = A^*$ y se cumplen las condiciones (3) del § 1 ó las condiciones equivalentes (véase el cap. VI, § 2, punto 3)

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0,$$

entonces

$$\rho = (1 - \xi)/(1 + \xi), \quad \xi = \gamma_1/\gamma_2;$$

2) si $A \neq A^*$ y se cumplen las condiciones (4) del § 1, ó las condiciones equivalentes (véase el cap. VI, § 4, punto 2)

$$\gamma_1 B \leqslant A, \quad (Ax, B^{-1}Ax) \leqslant \gamma_2 (Ax, x), \quad \gamma_1 > 0,$$

entonces

$$\rho = \sqrt{1 - \xi}, \quad \xi = \gamma_1/\gamma_2,$$

Señalemos que en comparación con los métodos del descenso más rápido y de los defectos mínimos en el método de las correcciones mínimas se exige invertir el operador B no una sola vez, sino dos, primeramente para calcular la corrección w_k, y después para calcular $B^{-1}Aw_k$.

Señalemos también, que si $A = A^*$, entonces el método de las correcciones mínimas posee la propiedad asintótica.

4. Método de los errores mínimos. Este método se aplica al igual que el método de los defectos mínimos, en el caso de cualquier operador A no autoconjugado y no degenerado.

El *método de los errores mínimos* se determina por la siguiente elección de los operadores B y D:

$$B = (A^*)^{-1}B_0, \quad D = B_0,$$

donde B_0 es un operador autoconjugado y definido positivo en H.

Sustituyendo el operador D elegido en la fórmula (3) para el parámetro iterativo τ_{k+1} y teniendo en cuenta, que $w_k = B^{-1}r_k = B_0^{-1}A^*r_k$, obtendremos la fórmula para τ_{k+1} en el método de los errores mínimos

$$\tau_{k+1} = \frac{(r_k, r_k)}{(Aw_k, r_k)}, \quad k = 0, 1, \ldots$$

La corrección w_k se halla de la ecuación $B_0 w_k = A^* r_k$.

En el caso de un esquema explícito ($B_0 = E$) la fórmula para τ_{k+1} tiene la forma

$$\tau_{k+1} = \frac{(r_k, r_k)}{(A^*r_k, A^*r_k)}, \quad k = 0, 1, \ldots$$

En el método de los errores mínimos se minimiza la norma del error en H_{B_0}. Para este método el operador $DB^{-1}A = A^*A$ es autoconjugado en H, y las condiciones (3) del § 1 toman la forma de las desigualdades (6). Del teorema 1 se deduce la estimación de la convergencia del método

$$\| z_n \|_{B_0} \leqslant \rho^n \| z_0 \|_{B_0}, \quad n \geqslant n_0(\varepsilon) = \ln \varepsilon / \ln \rho,$$

donde $\rho = (1 - \xi)/(1 + \xi)$, $\xi = \gamma_1/\gamma_2$, y γ_1, γ_2 están definidos en (6).

El método de los errores mínimos siempre posee la propiedad asintótica.

5. Ejemplo de aplicación de los métodos de dos capas. Para ilustrar la aplicación de los métodos gradientales de dos capas construidos examinaremos la resolución de un problema modelo por el método explícito del descenso más rápido. En calidad de ejemplo tomaremos el problema de Dirichlet de diferencias para la ecuación de Poisson sobre la red cuadrada $\bar{\omega} = \{x_{ij} = (ih, jh),\ 0 \leqslant i \leqslant N,\ 0 \leqslant j \leqslant N,\ h = 1/N\}$ en el cuadrado unidad

$$\Lambda u = u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2} = -\varphi, \quad x \in \omega, \quad u|_\gamma = g. \tag{7}$$

Introduzcamos el espacio H, que consiste de las funciones reticulares definidas sobre ω, con el producto escalar $(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h^2$.

Definiremos el operador A sobre H de la siguiente forma: $Ay = -\Lambda v$, $y \ni H$, donde $v(x) = y(x)$ para $x \in \omega$ y $v|_\gamma = 0$. Escribamos el problema (7) en forma de la ecuación operacional

$$Au = f, \tag{8}$$

donde f se diferencia de φ solamente en los nodos fronterizos

$$f = \varphi + \frac{\varphi_1}{h^2} + \frac{\varphi_2}{h^2},$$

$$\varphi_1 = \begin{cases} g(0, x_2), & x_1 = h, \\ 0, & 2h \leqslant x_1 \leqslant 1-2h, \\ g(1, x_2), & x_1 = 1-h, \end{cases} \qquad \varphi_2 = \begin{cases} g(x_1, 0), & x_2 = h, \\ 0, & 2h \leqslant x_2 \leqslant 1-2h, \\ g(x_1, 1), & x_2 = 1-h. \end{cases}$$

El operador A es autoconjugado y definido positivo en H. Por eso para resolver la ecuación (8) se puede aplicar el método del descenso más rápido. El esquema iterativo explícito tiene la forma

$$\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f \quad \text{o} \quad y_{k+1} = y_k - \tau_{k+1} r_k, \quad k = 0, 1, \ldots,$$

y los parámetros iterativos τ_k se encuentran por la fórmula

$$\tau_{k+1} = \frac{(r_k, r_k)}{(Ar_k, r_k)}, \quad r_k = Ay_k - f, \quad k = 0, 1, \ldots$$

Citemos las fórmulas de cálculo y contemos el número de operaciones aritméticas, gastadas en una iteración. Teniendo en cuenta la definición del operador A y del miembro derecho f, las fórmulas de cálculo se pueden escribir en la forma siguiente:

1) $r_k(x_{ij}) = -(y_k)_{\bar{x}_1 x_1} - (y_k)_{\bar{x}_2 x_2} - \varphi(x_{ij}), \quad 1 \leqslant i, j \leqslant N-1, \quad y_k|_\gamma = g;$

2) $$(r_k, r_k) = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} r_k^2(x_{ij})\, h^2,$$

$$(Ar_k, r_k) = -\sum_{i=1}^{N-1} \sum_{j=1}^{N-1} r_k(x_{ij}) [(r_k)_{\bar{x}_1 x_1} + (r_k)_{\bar{x}_2 x_2}]\, h^2,$$

$$r_k|_\gamma = 0, \quad \tau_{k+1} = \frac{(r_k, r_k)}{(Ar_k, r_k)};$$

3) $y_{k+1}(x_{ij}) = y_k(x_{ij}) - \tau_{k+1} r_k(x_{ij}), \quad 1 \leqslant i, j \leqslant N-1.$

La aproximación inicial y_0 es función reticular arbitraria en ω, que toma sobre γ valores prefijados $y_0|_\gamma = g$.

Contemos el número de operaciones aritméticas. Si realizamos el cálculo de las derivadas en diferencias por la fórmula

$$u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2} = \frac{1}{h^2}(u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + U_{i,j-1} - 4u_{ij}),$$

entonces para calcular r_k se exige $6(N-1)^2$ adiciones y $2(N-1)^2$ multiplicaciones y divisiones. Para calcular (r_k, r_k) se exige $(N-1)^2$ adiciones y $(N-1)^2$ multiplicaciones, para (Ar_k, r_k) — $6(N-1)^2$ adiciones y $2(N-1)^2$ multiplicaciones, y para y_{k+1} — $(N-1)^2$ adiciones y $(N-1)^2$ multiplicaciones. En total hay que efectuar $14(N-1)^2$ adiciones y $6(N-1)^2$ multiplicaciones y divisiones. Exactamente la mitad de este número total de operaciones necesita para calcular los productos escalares, o sea para el cálculo del parámetro iterativo τ_{k+1}. Por consiguiente, un paso del método del descenso más rápido es aproximadamente dos veces más laborioso en comparación con un paso del método de iteración simple o del método de Chebishev, donde los parámetros τ_{k+1} son conocidos a priori. Para los métodos implícitos esta diferencia será menor, puesto que para el cálculo de los productos escalares se exige el mismo número de operaciones que para el método explícito, y al número total de operaciones se añaden las operaciones aritméticas gastadas en la inversión del operador B.

Determinemos ahora el número total de operaciones aritméticas $Q(\varepsilon)$, que es necesario cumplir, para obtener una exactitud relativa ε. Para esto es necesario estimar el número de iteraciones $n_0(\varepsilon)$. En el punto 1 fue obtenida la estimación siguiente

$$n_0(\varepsilon) = \frac{\ln \varepsilon}{\ln \rho}, \quad \rho = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

donde γ_1 y γ_2 en el caso de un esquema explícito son las cotas del operador A: $\gamma_1 E \leqslant A \leqslant \gamma_2 E$.

Para el ejemplo examinado γ_1 y γ_2 coinciden con los valores propios mínimo δ y máximo Δ del operador de Laplace de diferencias Λ. Es conocido, que

$$\delta = \frac{8}{h^2}\operatorname{sen}^2\frac{\pi h}{2}, \quad \Delta = \frac{8}{h^2}\cos^2\frac{\pi h}{2}.$$

Por eso

$$\rho = \frac{1-\xi}{1+\xi} = 1 - 2\operatorname{sen}^2\frac{\pi h}{2}, \quad \xi = \frac{\delta}{\Delta} = \operatorname{tg}^2\frac{\pi h}{2},$$

y, por lo tanto, si $h \ll 1$, entonces

$$n_0(\varepsilon) \approx \frac{2\ln\frac{1}{\varepsilon}}{\pi^2 h^2} \approx 0{,}2N^2 \ln\frac{1}{\varepsilon}.$$

Si consideramos equivalentes las operaciones de adición, multiplicación y división, entonces en una iteración se gastan aproximadamente $20N^2$ operaciones. Por eso para el número total de operaciones aritméticas será válida la estimación $Q(\varepsilon) \approx 4N^4 \ln\frac{1}{\varepsilon}$.

§ 3. Métodos de tres capas de las direcciones conjugadas

1. Planteamiento del problema sobre la elección de los parámetros iterativos. Estimación de la velocidad de convergencia. Para encontrar una solución aproximada de la ecuación operacional lineal

$$Au = f \tag{1}$$

con el operador no degenerado A en el § 1 nosotros examinamos los métodos iterativos de dos capas de tipo variacional. El esquema iterativo para estos métodos tiene la forma

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \quad y_0 \in H, \tag{2}$$

y los parámetros iterativos τ_{k+1} se eligen de la condición de mínimo de la norma del error z_{k+1} en el espacio energético H_D. Recordemos que sobre la sucesión y_k construida según la fórmula (2), se realiza la minimización por pasos del funcional $I(y) = (D(y-u), y-u)$, cuyo mínimo se alcanza en la solución de la ecuación (1), es decir, si $y = u$.

Sin embargo, esta estrategia de minimización local no es óptima, puesto que en fin de cuentas a nosotros nos interesa el mínimo global del funcional $I(y)$, y, si está prefijado un cierto valor de este funcional, debemos alcanzar el mínimo de iteraciones. La minimización local en cada paso iterativo conduce a la resolución de este problema pero no por el camino más corto.

Es natural tratar de elegir el parámetro τ_k de la condición de mínimo de la norma del error z_n en H_D inmediatamente después de n pasos, o sea al pasar de y_0 a y_n. Con una situa-

ción análoga nosotros ya nos encontramos en el capítulo VI durante el estudio del método de Chebishev y del método de iteración simple. Resultó que converge más rápidamente aquel método cuyos parámetros iterativos se eligen de la condición de mínimo de la norma del operador de resolución, y no del operador de transición de una iteración a otra. Esta propiedad tiene lugar también para los métodos iterativos de tipo variacional. Será mostrado, que los métodos iterativos examinados en este párrafo, para los cuales los parámetros τ_k se eligen de la condición indicada más arriba, convergen considerablemente más rápido que los métodos gradientales de dos capas. Además, en el caso de un espacio H de dimensión finita estos métodos son los métodos de finitas iteraciones para toda aproximación inicial, es decir la solución exacta de la ecuación (1) puede ser obtenida después de un número finito de iteraciones.

Pasemos a la construcción del *método de las direcciones conjugadas*. Supondremos que el operador $DB^{-1}A$ es autoconjugado y definido positivo en H. Cumplimos n iteraciones según el esquema (2). Pasando del problema para el error $z_k = y_k - u$ al problema para $x_k = D^{1/2}z_k$, obtendremos, como antes,

$$x_{k+1} = S_{k+1}x_k, \quad k = 0, 1, \ldots, n-1,$$
$$S_k = E - \tau_k C, \quad C = D^{1/2}B^{-1}AD^{-1/2}.$$

De aquí hallaremos

$$x_n = T_n x_0, \quad T_n = \prod_{j=1}^{n} (E - \tau_j C). \tag{3}$$

El operador de resolución T_n representa un polinomio operacional de grado n respecto al operador C con coeficientes que dependen de los parámetros $\tau_1, \tau_2, \ldots, \tau_n$

$$T_n = P_n(C) = E + \sum_{j=1}^{n} a_j^{(n)} C^j, \quad a_n^{(n)} \neq 0. \tag{4}$$

En virtud de la igualdad $\| x_n \| = \| z_n \|_D$ el problema, planteado más arriba, sobre la elección de los parámetros iterativos τ_k se formula de la siguiente manera: entre todos los polinomios del tipo (4) elegir aquel, para el cual la norma de $x_n = P_n(C)\, x_0$ sea mínima, en otras palabras —elegir los coeficientes $a_1^{(n)}, a_2^{(n)}, \ldots, a_n^{(n)}$ del polinomio $P_n(C)$ de la condición de mínimo de la norma de x_n en H.

Este problema será resuelto en el punto siguiente, y primeramente obtendremos una estimación de la velocidad de convergencia del método de las direcciones conjugadas, construido a base del principio de elección de los parámetros formulado más arriba. Obtendremos esta estimación utilizando la información a priori sobre los operadores del esquema en forma de las constantes γ_1 y γ_2 de equivalencia energética de los operadores autoconjugados D y $DB^{-1}A$:

$$\gamma_1 D \leqslant DB^{-1}A \leqslant \gamma_2 D, \quad \gamma_1 > 0, \; DB^{-1}A = (DB^{-1}A)^*. \quad (5)$$

Supongamos que $P_n(C)$ es el polinomio buscado. Entonces de (3) y (4) se deduce una estimación de x_n:

$$\| x_n \| = \| P_n(c)\, x_0 \| = \| \min_{\{Q_n\}} \| Q_n(C)\, x_0 \| \leqslant$$

$$\leqslant \min_{\{Q_n\}} \| Q_n(C) \| \, \| x_0 \|,$$

donde el mínimo se busca entre los polinomios $Q_n(C)$ normados en virtud de (4) por la condición $Q_n(0) = E$.

Expresemos el mínimo de la norma del polinomio $Q_n(C)$. De (5) se deduce que el operador $C = D^{-1/2}(DB^{-1}A)\,D^{-1/2}$ es autoconjugado en H, mientras que γ_1, γ_2 son sus cotas: $C = C^*$, $\gamma_1 E \leqslant C \leqslant \gamma_2 E$, $\gamma_1 > 0$. Por eso tiene lugar la estimación

$$\min_{\{Q_n\}} \| Q_n(C) \| \leqslant \min_{\{Q_n\}} \max_{\gamma_1 \leqslant t \leqslant \gamma_2} | Q_n(t) |.$$

De los resultados del § 2 del cap. VI se desprende que el problema acerca de la construcción de un polinomio normado por la condición $Q_n(0) = 1$ cuyo módulo máximo, sobre el segmento $[\gamma_1, \gamma_2]$ sea mínimo, lo resuelve el polinomio de Chebishev de 1er género, para el cual

$$\max_{\gamma_1 \leqslant t \leqslant \gamma_2} | Q_n(t) | = q_n, \quad q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}},$$

$$\rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Por consiguiente, para x_n tiene lugar la estimación $\| x_n \| \leqslant q_n \| x_0 \|$.

Así, está demostrado

TEOREMA 2. *Si son* satisfechas *las condiciones* (5), *entonces el método iterativo de las direcciones conjugadas converge en* H_D, *y con cualquier n para el error* z_n *es válida la estimación*

$\| z_n \|_D \leqslant q_n \| z_0 \|_D$. *En este caso la estimación para el número de iteraciones tiene la forma*

$$n \geqslant n_0(\varepsilon) = \ln(0{,}5\varepsilon)/\ln \rho_1,$$

donde $\rho_1 = (1 - \sqrt{\xi})/(1 + \sqrt{\xi})$, $\xi = \gamma_1/\gamma_2$.

2. Fórmulas para los parámetros iterativos. Esquema iterativo de tres capas. Pasemos ahora a la construcción del polinomio $P_n(C)$. Utilizando (3) y (4), calculemos la norma de x_n:

$$\| x_n \|^2 = (P_n(C) x_0,\ P_n(C) x_0) = \| x_0 \|^2 + \\ + 2 \sum_{j=1}^{n} a_j^{(n)} (C^j x_0,\ x_0) + \sum_{j=1}^{n} \sum_{i=1}^{n} a_j^{(n)} a_i^{(n)} (C^j x_0,\ C^i x_0).$$

La norma de x_n es una función de los parámetros $a_1^{(n)}$, $a_2^{(n)}$, ..., $a_n^{(n)}$. Igualando a cero las derivadas parciales de $\| x_n \|^2$ respecto a $a_j^{(n)}$

$$\frac{\partial \| x_n \|^2}{\partial a_j^{(n)}} = 2 \sum_{i=1}^{n} a_i^{(n)} (C^j x_0,\ C^i x_0) + 2 (C^j x_0,\ x_0),\ j = 1, 2, \ldots, n,$$

obtendremos el sistema de ecuaciones algebraicas lineales

$$\sum_{i=1}^{n} a_i^{(n)} (C^j x_0,\ C^i x_0) + (C^j x_0,\ x_0) = 0,\ j = 1, 2, \ldots, n. \qquad (6)$$

Para un operador C autoconjugado y definido positivo en H el sistema (6) da la condición de mínimo de la norma de x_n en H.

De esta forma, el problema de la construcción de un polinomio óptimo $P_n(C)$ en principio está resuelto. Hallemos los coeficientes $a_1^{(n)}$, $a_2^{(n)}$, ..., $a_n^{(n)}$ del polinomio resolviendo el sistema (6). Pero primeramente construyamos las fórmulas para calcular la aproximación iterativa y_n. El primer camino consiste en el empleo del esquema iterativo (2). Sin embargo, para esto hay que hallar las raíces del polinomio $P_n(t)$ y después en calidad de τ_k tomar los valores inversos a las raíces. Este método no es económico.

El segundo método consiste en utilizar los coeficientes del polinomio para calcular y_n. De (3), (4) y del cambio $x_k = D^{1/2} z_k$, donde $z_k = y_k - u$, obtenemos

$$y_n - u = D^{-1/2} P_n(C) D^{1/2} (y_0 - u). \qquad (7)$$

Utilizando (4) y la igualdad $D^{-1/2}C^jD^{1/2} = (B^{-1}A)^j$, hallaremos

$$D^{-1/2}P_n(C)\,D^{1/2} = E + \sum_{j=1}^{n} a_j^{(n)} (B^{-1}A)^j.$$

Sustituyendo esta igualdad en (7), obtendremos

$$y_n = y_0 + \sum_{j=1}^{n} a_j^{(n)} (B^{-1}A)^j (y_0 - u) =$$

$$= y_0 + \sum_{j=1}^{n} a_j^{(n)} (B^{-1}A)^{j-1} w_0, \quad (8)$$

donde w_0 es la corrección, $w_0 = B^{-1}A\,(y_0 - u) = B^{-1}r_0$, $r_0 = Ay_0 - f$.

Este camino tampoco es óptimo. Para cada n se tiene que resolver de nuevo el sistema (6).

Ahora mostraremos que la sucesión $y_1, y_2, \ldots, y_k, \ldots$, construida de acuerdo con (6) y (8) para $n = 1, 2, \ldots$, puede ser hallada por el siguiente esquema de tres capas: $By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)\,y_k + (1 - \alpha_{k+1})\,By_{k-1} +$

$$+ \alpha_{k+1}\tau_{k+1}f, \quad k = 1, 2, \ldots, \quad (9)$$

$$By_1 = (B - \tau_1 A)\,y_0 + \tau_1 f, \quad y_0 \in H.$$

Para esto es necesario indicar la colección de parámetros $\{\tau_k\}$ y $\{\alpha_k\}$, para la cual la norma del error equivalente x_k sea mínima para cualquier k. Realmente, de la ecuación para el error x_k en el caso del esquema (9)

$$x_{k+1} = \alpha_{k+1}(E - \tau_{k+1}C)\,x_k + (1 - \alpha_{k+1})\,x_{k-1},$$

$$k = 1, 2, \ldots, \quad (10)$$

$$x_1 = (E - \tau_1 C)\,x_0,$$

obtendremos, que $x_k = P_k(C)\,x_0$, donde el polinomio $P_k(C)$ tiene la forma (4) ($n = k$). Por esto, si los parámetros $\{\tau_k\}$ y $\{\alpha_k\}$ en (9) serán elegidos de manera tal, que para cualquier $n = 1, 2, \ldots$ se cumplan las condiciones (6), entonces las aproximaciones iterativas y_n construidas de acuerdo con (9) coincidirán con las aproximaciones, construidas mediante las fórmulas (6) y (8) para cualquier n.

Construyamos el conjunto buscado de parámetros $\{\tau_k\}$ y $\{\alpha_k\}$. Para esto necesitaremos

LEMA. *Las condiciones necesarias y suficientes de mínimo de la norma de x_n en H para cualquier $n \geqslant 1$ son*

$$(Cx_j,\ x_n) = 0, \quad j = 0,\ 1,\ \ldots,\ n-1. \tag{11}$$

En efecto, de (4) y (6) se sigue que las condiciones (6) que son las condiciones de mínimo de la norma de x_n, son equivalentes a las siguientes:

$$(C^j x_0,\ x_n) = 0, \quad j = 1,\ 2,\ \ldots,\ n, \tag{12}$$

para cualquier $n = 1,\ 2,\ \ldots$. De aquí obtenemos para $j \leqslant n-1$

$$(Cx_0,\ x_n) + \sum_{i=2}^{j+1} a_{i-1}^{(j)} (C^i x_0,\ x_n) = (Cx_j,\ x_n) = 0,$$

es decir las condiciones (11) son necesarias.

Demostremos ahora la suficiencia de las condiciones (11). Supongamos que se cumplen las condiciones (11). Mostremos que entonces también se cumplen las condiciones (12). De (11) para $j = 0$ obtenemos que las igualdades (12) son ciertas para $j = 1$. Demostraremos por inducción la validez de (12) para $j \geqslant 2$. Supongamos que para $j \leqslant k$ ya están cumplidas las condiciones (12), es decir, $(C^j x_0,\ x_n) = 0$, $j = 1,\ 2,\ \ldots$ $\ldots,\ k$. Mostremos que ellas también son cumplidas para $j = k+1$, si se cumplen las condiciones (11).

Efectivamente, de (11 para $j = k$ obtenemos

$$0 = (Cx_k,\ x_n) = (CP_k(C)\,x_0,\ x_n) = (Cx_0,\ x_n) + \\ + \sum_{j=1}^{k} a_j^{(k)} (C^{j+1} x_0,\ x_n) = a_k^{(k)} (C^{k+1} x_0,\ x_n).$$

Por lo tanto, $(C^{k+1} x_0,\ x_n) = 0$. El lema está demostrado.

Utilicemos ahora el lema para construir la colección de los parámetros $\{\tau_k\}$ y $\{\alpha_k\}$ para el esquema (9). Para abreviar los razonamientos consideraremos que y_1 en el esquema (9) se halla por la fórmula general (9) para $\alpha_1 = 1$.

Examinemos el esquema (10). Ya que x_1 se encuentra por un esquema de dos capas, entonces del § 1 se deduce la elección del parámetro óptimo τ_1 por la fórmula

$$\tau_1 = \frac{(Cx_0,\ x_0)}{(Cx_0,\ Cx_0)},$$

Realizaremos gradualmente la construcción de los parámetros $\tau_2,\ \tau_3,\ \ldots$ y $\alpha_2,\ \alpha_3,\ \ldots$ Sean iterativos los parámetros $\tau_1,\ \tau_2,\ \ldots,\ \tau_k$ y $\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_k$ ya elegidos de

manera óptima. Puesto que estos parámetros determinan las aproximaciones $y_1, y_2, \ldots, y_k$, entonces del lema se deduce que se cumplen las condiciones

$$(Cx_j, x_i) = 0, \quad j = 0, 1, \ldots, i-1, \ i = 1, 2, \ldots, k. \quad (13)$$

Elijamos ahora los parámetros τ_{k+1} y α_{k+1}, que determinan la aproximación y_{k+1}. Del lema se sigue que la norma de x_{k+1} será mínima si están cumplidas las condiciones

$$(Cx_j, x_{k+1}) = 0, \quad j = 0, 1, \ldots, k. \quad (14)$$

De estas condiciones hallaremos los parámetros τ_{k+1} y α_{k+1}. Mostremos primeramente que de (13) se deduce el cumplimiento de las condiciones (14) para $j \leqslant k-2$, y después de las dos condiciones restantes de (14) para $j = k-1$ y $j = k$ obtendremos las fórmulas para τ_{k+1} y α_{k+1}.

Así pues, sea $j \leqslant k-2$. De (10) y (13) hallaremos

$$(x_{k+1}, Cx_j) = \alpha_{k+1}(x_k, Cx_j) - \alpha_{k+1}\tau_{k+1}(Cx_k, Cx_j) + (1-\alpha_{k+1})(x_{k-1}, Cx_j) = -\alpha_{k+1}\tau_{k+1}(Cx_k, Cx_j).$$

Mostremos que $(Cx_k, Cx_j) = 0$ para $j \leqslant k-2$. En realidad, de (10) para $k = j$ obtenemos

$$Cx_j = \frac{1}{\tau_{j+1}} x_j - \frac{1}{\tau_{j+1}\alpha_{j+1}}[x_{j+1} - (1-\alpha_{j+1})x_{j-1}], \ j \geqslant 0. \quad (15)$$

Utilizando la autoconjugación del operador C y las condiciones (13), de aquí obtendremos para $j \leqslant k-2$

$$(Cx_k, Cx_j) = \frac{1}{\tau_{j+1}}(Cx_j, x_k) - \frac{1}{\tau_{j+1}\alpha_{j+1}}[(Cx_{j+1}, x_k) - (1-\alpha_{j+1})(Cx_{j-1}, x_k)] = 0.$$

Por consiguiente, $(x_{k+1}, Cx_j) = 0$ para $j \leqslant k-2$.

Hallemos ahora τ_{k+1} y α_{k+1}. Poniendo en (14) $j = k-1$ y $j = k$, obtenemos de (10) y (13)

$$0 = (Cx_{k-1}, x_{k+1}) = -\alpha_{k+1}\tau_{k+1}(Cx_k, Cx_{k-1}) + (1-\alpha_{k+1})(Cx_{k-1}, x_{k-1}),$$

$$0 = (Cx_k, x_{k+1}) = \alpha_{k+1}[(Cx_k, x_k) - \tau_{k+1}(Cx_k, Cx_k)]. \quad (16)$$

De la segunda ecuación hallamos inmediatamente el parámetro τ_{k+1}:

$$\tau_{k+1} = \frac{(Cx_k,\ x_k)}{(Cx_k,\ Cx_k)}. \tag{17}$$

De la primera ecuación eliminaremos la expresión $(Cx_k,\ Cx_{k-1})$. Para eso podremos $j = k - 1$ en (15) y multipliquemos escalarmente los miembros primero y segundo de (15) por Cx_k.

Teniendo en cuenta la autoconjugación del operador C de la condición (13), obtenemos

$$(Cx_k,\ Cx_{k-1}) = \frac{1}{\tau_k}(Cx_{k-1},\ x_k) - \frac{1}{\tau_k\alpha_k}(Cx_k,\ x_k) + \\ + \frac{1-\alpha_k}{\tau_k\alpha_k}(Cx_{k-2},\ x_k) = -\frac{1}{\tau_k\alpha_k}(Cx_k,\ x_k).$$

Sustituyendo esta expresión en (16), obtendremos

$$\frac{\alpha_{k+1}\tau_{k+1}}{\alpha_k\tau_k}\frac{(Cx_k,\ x_k)}{(Cx_{k-1},\ x_{k-1})} + (1-\alpha_{k+1}) = 0.$$

De esta igualdad hallaremos la fórmula recurrente para el parámetro α_{k+1}:

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k}\frac{(Cx_k,\ x_k)}{(Cx_{k-1},\ x_{k-1})}\frac{1}{\alpha_k}\right)^{-1}. \tag{18}$$

Así, suponiendo que los parámetros iterativos $\tau_1, \tau_2, \ldots$ $\ldots, \tau_k$ y $\alpha_1, \alpha_2, \ldots, \alpha_k$ fueron elegidos anteriormente, nosotros obtenemos fórmulas para los parámetros τ_{k+1} y α_{k+1}. Puesto que $\alpha_1 = 1$ y $\tau_1 = \frac{(Cx_0, Cx_0)}{(Cx_0, Cx_0)}$, entonces las fórmulas (17) y (18) determinan los parámetros τ_{k+1} y α_{k+1} para cualquier k.

Sustituyendo $x_k = D^{1/2}z_k$ en (17) y (18) y teniendo en cuenta que

$$C = D^{-1/2}(DB^{-1}A)D^{-1/2} \text{ y } Az_k = r_k,\ B^{-1}r_k = w_k,$$

obtendremos las fórmulas siguientes para los parámetros iterativos τ_{k+1} y α_{k+1}:

$$\tau_{k+1} = \frac{(Dw_k,\ z_k)}{(Dw_k,\ w_k)},\ k = 0,\ 1,\ \ldots, \tag{19}$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k}\frac{(Dw_k,\ z_k)}{(Dw_{k-1},\ z_{k-1})}\frac{1}{\alpha_k}\right)^{-1} \tag{20}$$

$k = 1,\ 2,\ \ldots,\ \alpha_1 = 1.$

Así, el método de las direcciones conjugadas se describe por el esquema de tres capas (9), para el cual los parámetros

iterativos τ_{k+1} y α_{k+1} se eligen por las fórmulas (19) y (20). Para este método es válido el teorema 2 demostrado anteriormente.

De las fórmulas (19) y (20) se deduce que los parámetros iterativos τ_{k+1} en el método de las direcciones conjugadas y en los métodos gradientales de dos capas se seleccionan por las mismas fórmulas, y para calcular los parámetros α_{k+1} no es necesario calcular ningunos productos escalares complementarios. Por esto al calcular los parámetros iterativos mediante los métodos de dos y tres capas de tipo variacional se gasta prácticamente igual número de operaciones aritméticas. A su vez de los teoremas 1 y 2 se sigue que los métodos de las direcciones conjugadas convergen esencialmente más rápido que los métodos gradientales.

Mostremos ahora, que si H es un espacio de dimensión finita ($H = H_N$), entonces los métodos de las direcciones conjugadas convergen durante un número finito de iteraciones, no superior a la dimensión del espacio. Realmente, del lema se deduce que para los errores equivalentes x_k del método de las direcciones conjugadas deben cumplirse las igualdades $(Cx_j, x_n) = (x_j, x_n)_C = 0, \quad j = 0, 1, \ldots, n-1$. Por consiguiente, el sistema de vectores $x_0, x_1, \ldots, x_n$ para cualquier n debe ser un sistema ortogonal en H_C. Y como en H_N no se puede construir más de N vectores ortogonales, entonces de aquí se desprende que $x_N = 0$ y $z_N = y_N - u = 0$. De esta forma, en la clase de las aproximaciones iniciales arbitrarias y_0 los métodos de las direcciones conjugadas convergen después de N iteraciones a la solución exacta de la ecuación (1).

Para aproximaciones iniciales especiales y_0 estos métodos convergen en un número menor de iteraciones. En efecto, sea y_0 tal, que en el desarrollo de x_0 según las funciones propias del operador C estén presentes $N_0 < N$ funciones, es decir, x_0 pertenece al subespacio H_{N_0} invariante con respecto al operador C. Entonces es evidete, que todos los $x_k \in H_{N_0}$. Por eso en este caso el proceso iterativo termina por N_0 iteraciones.

De lo dicho más arriba se deduce que la estimación de la convergencia del método obtenido en el teorema 2, es muy aproximada y la igualdad $\| z_n \|_D = q_n \| z_0 \|_D$ nunca se alcanza. Se puede construir un ejemplo de ecuación (1) e indicar para cualquier $n < N$ una aproximación inicial y_0 tal que se cumplirá la igualdad indicada.

3. Variantes de las fórmulas de cálculo. Citemos ahora algunos procedimientos de realización de los métodos de tres capas de direcciones conjugadas. De (9), (19) y (20) obtenemos el siguiente algoritmo:

1) por y_0 dado se calcula el defecto $r_0 = Ay_0 - f$;

2) se resuelve la ecuación para la corrección $Bw_0 = r_0$;

3) se calcula el parámetro τ_1 por la fórmula (19);

4) la aproximación y_1 se encuentra por la fórmula $y_1 = y_0 - \tau_1 w_0$.

Después para $k = 1, 2, \ldots$ sucesivamente se cumplen las siguientes operaciones:

5) se calcula el defecto $r_k = Ay_k - f$ y se resuelve la ecuación para la corrección $Bw_k = r_k$;

6) por las fórmulas (19) y (20) se calculan los parámetros τ_{k+1} y α_{k+1};

7) la aproximación y_{k+1} se encuentra por la fórmula

$$y_{k+1} = \alpha_{k+1}y_k + (1 - \alpha_{k+1})\, y_{k-1} - \alpha_{k+1}\tau_{k+1}w_k.$$

De esta forma, en el algoritmo descrito para encontrar y_{k+1} se utilizan y_{k-1}, y_k y w_k los cuales es necesario memorar. Más abajo será indicado el tipo de las fórmulas (19) y (20) para algunas elecciones concretas del operador D. Aquí nos limitaremos a la observación de que en estas fórmulas, además de la magnitud w_k memorado, puede entrar el defecto r_k, el cual nosotros no memoramos. Para su cálculo se puede utilizar o bien la igualdad $r_k = Bw_k$, si el cálculo del valor de Bw_k no es un proceso laborioso, o la definición del defecto $r_k = Ay_k - f$.

En la práctica se encuentran también otros algoritmos de realización del método de las direcciones congugadas. Citemos uno de ellos. Para eso interpretaremos el esquema (9) como un esquema con corrección. De (9) obtenemos

$$y_{k+1} = y_k - a_{k+1}s_k, \quad s_{k+1} = w_{k+1} + b_{k+1}s_k,$$
$$k = 0, 1 \ldots, s_0 = w_0, \tag{21}$$

donde $w_k = B^{-1}r_k$, $r_k = Ay_k - f$, y los parámetros a_{k+1} y b_k están relacionados con α_{k+1} y τ_{k+1} por las fórmulas siguientes:

$$a_{k+1} = \alpha_{k+1}\tau_{k+1}, \quad b_k = (\alpha_{k+1} - 1)\,\alpha_k\tau_k/(\alpha_{k+1}\tau_{k+1}).$$

Obtengamos las expresiones para a_{k+1} y b_k. De (19) y (20) hallaremos

$$b_k = (Dw_k, z_k)/(Dw_{k-1}, z_{k-1}), \quad k = 1,2, \ldots \tag{22}$$

Para a_{k+1} de estas fórmulas es fácil obtener relaciones recurrentes, sin embargo se puede hallar una expresión explícita para a_{k+1}:

$$a_{k+1} = \frac{(Cx_k, x_k)}{(p_k, p_k)} = \frac{(Dw_k, z_k)}{(Ds_k, s_k)}, \quad k = 0, 1, \ldots \tag{23}$$

Las fórmulas (21), (22) y (23) describen el segundo algoritmo del método de las direcciones conjugadas. Aquí los cálculos se realizan en el orden siguiente:

1) por y_0 dado se calcula el defecto $r_0 = Ay_0 - f$, se resuelve la ecuación $Bw_0 = r_0$ para la corrección w_0, y se pone $s_0 = w_0$;

2) por la fórmula (23) se halla el parámetro a_1 y se calcula $y_1 = y_0 - a_1 s_0$. Luego para $k = 1, 2, \ldots$ sucesivamente se efectúan las operaciones;

3) se calcula el defecto $r_k = Ay_k - f$ y se resuelve la ecuación para la corrección $Bw_k = r_k$;

4) por la fórmula (22) se calcula el parámetro b_k y se encuentra s_k por la fórmula $s_k = w_k + b_k s_{k-1}$;

5) por la fórmula (23) se determina el parámetro a_{k+1} y la aproximación y_{k+1} se calcula según la fórmula

$$y_{k+1} = y_k - a_{k+1} s_k.$$

Señalemos que en el algoritmo citado es necesario conservar y_k, w_k y s_k, es decir el mismo volumen de información intermedia que en el primer algoritmo.

§ 4. Ejemplos de métodos de tres capas

1. Casos particulares de los métodos de las direcciones conjugadas. En el § 3 fueron construidos los métodos iterativos de tres capas de las direcciones conjugadas, utilizados para resolver la ecuación lineal

$$Au = f. \tag{1}$$

Las aproximaciones iterativas se calculan por el esquema de tres capas

$$By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)\, y_k + (1 - \alpha_{k+1})\, By_{k-1} + \\ + \alpha_{k+1}\tau_{k+1} f, \quad k = 1, 2, \ldots, \tag{2}$$

$$By_1 = (B - \tau_1 A)\, y_0 + \tau_1 f, \quad y_0 \in H,$$

y los parámetros iterativos α_{k+1} y τ_{k+1} se encuentran por las fórmulas

$$\tau_{k+1} = \frac{(Dw_k,\ z_k)}{(Dw_k,\ w_k)}, \quad k = 0, 1, \ldots, \tag{3}$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k} \frac{(Dw_k,\ z_k)}{(Dw_{k-1},\ z_{k-1})} \cdot \frac{1}{\alpha_k}\right)^{-1}, \\ k = 1, 2, \ldots, \ \alpha_1 = 1,$$

donde $w_k = B^{-1} r_k$ es la corrección, $r_k = Ay_k - f$ es el defecto y $z_k = y_k - u$ el error.

La elección de los parámetros α_k y τ_k según las fórmulas (3) en el caso del operador autoconjugado y definido positivo $DB^{-1}A$ asegura el mínimo para cualquier n de la norma del error z_n en H_D al pasar de y_0 a y_n.

Examinemos ahora casos particulares de los métodos de las direcciones conjugadas, determinados mediante la elección del operador D. En el § 2 fueron examinados cuatro ejemplos de métodos gradientales de dos capas. A cada uno de estos métodos de dos capas le corresponde un determinado método de tres capas de direcciones conjugadas. Nosotros enumeraremos estos métodos, indicando las condiciones sobre los operadores A y B, las cuales garantizan la autoconjugación del operador $DB^{-1}A$. Para estos métodos tiene lugar el teorema 2, y el tipo de las desigualdades que determinan las constantes γ_1 y γ_2, será indicado en la descripción del método correspondiente.

1) *Método de los gradientes conjugados.*
Operador D: $D = A$.
Condiciones: $\gamma_1 B \leqslant A \leqslant \gamma_2 B$, $\gamma_1 > 0$, $A = A^* > 0$, $B = B^* > 0$.
Las fórmulas para los parámetros iterativos:

$$\tau_{k+1} = \frac{(r_k, w_k)}{(Aw_k, w_k)},$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k} \frac{(r_k, w_k)}{(r_{k-1}, w_{k-1})} \frac{1}{\alpha_k}\right)^{-1}.$$

2) *Método de los defectos conjugados.*
Operador D: $D = A^*A$.
Condiciones: $\gamma_1 (Bx, Bx) \leqslant (Ax, Bx) \leqslant \gamma_2 (Bx, Bx)$, $\gamma_1 > 0$, $B^*A = A^*B$.
Si se cumplen las suposiciones $A = A^* > 0$, $B = B^* > 0$, $AB = BA$, entonces las condiciones tienen la forma $\gamma_1 B \leqslant A \leqslant \gamma_2 B$, $\gamma_1 > 0$.
Las fórmulas para los parámetros iterativos:

$$\tau_{k+1} = \frac{(Aw_k, r_k)}{(Aw_k, Aw_k)},$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_k}{\tau_k} \frac{(Aw_k, r_k)}{(Aw_{k-1}, r_{k-1})} \cdot \frac{1}{\alpha_k}\right)^{-1}.$$

3) *Método de las correcciones conjugadas.*
Operador D: $D = AB^{-1}A$.
Condiciones: $\gamma_1 B \leqslant A \leqslant \gamma_2 B$, $\gamma_1 > 0$, $A = A^* > 0$, $B = B^* > 0$.
Fórmulas para los parámetros iterativos:

$$\tau_{k+1} = \frac{(Aw_k, w_k)}{(B^{-1}Aw_k, Aw_k)},$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k} \frac{(Aw_k, w_k)}{(Aw_{k-1}, w_{k-1})} \cdot \frac{1}{\alpha_k}\right)^{-1}.$$

4) *Método de los errores conjugados.*
Operador D: $D = B_0$.
Condiciones: $B = (A^*)^{-1}B_0$, $\gamma_1 B_0 \leqslant A^*A \leqslant \gamma_2 B_0$, $B_0 = B_0^* > 0$.
Fórmulas para los parámetros iterativos:

$$\tau_{k+1} = \frac{(r_k, r_k)}{(Aw_k, r_k)},$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k}\frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})}\cdot\frac{1}{\alpha_k}\right)^{-1}.$$

2. Métodos de tres capas localmente óptimos. Regresemos ahora al procedimiento de construcción de los parámetros iterativos α_{k+1} y τ_{k+1} examinado en el § 3 para el esquema de tres capas del método de las direcciones conjugadas. Recordemos que los parámetros α_{k+1} y τ_{k+1} fueron elegido de las condiciones $(Cx_{k-1}, x_{k+1}) = 0$ y $(Cx_k, x_{k+1}) = 0$ bajo la suposición de que las aproximaciones iterativas $y_1, y_2, \ldots, y_k$ aseguran el cumplimiento de las condiciones

$$(Cx_j, x_i) = 0,\ j = 0, 1, \ldots, i-1,\ i = 1, 2, \ldots, k. \quad (4)$$

Para un proceso computacional ideal se cumplen las condiciones (4), por eso la elección de los parámetros α_{k+1} y τ_{k+1} por las fórmulas obtenidas en el § 3 realmente garantiza el mínimo de la norma del error z_{k+1} en H_D al pasar de y_0 a y_{k+1}. En un proceso computacional real, el cual tiene en cuenta la presencia de errores de redondeo, las aproximaciones iteradas $y_1, y_2, \ldots, y_k$ no serán calculadas exactamente y, por consiguiente, no se cumplirán las condiciones (4). En una serie de casos esto puede conducir a la disminución de la velocidad de convergencia del método, y en algunas ocasiones a su divergencia.

Construyamos ahora una modificación del método de las direcciones conjugadas, la cual no posee la insuficencia indicada. Para resolver aproximadamente la ecuación $Au = f$ examinemos el esquema iterativo

$$By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)y_k + (1 - \alpha_{k+1})By_{k-1} + \alpha_{k+1}\tau_{k+1}f,$$
$$k = 1, 2, \ldots \quad (5)$$

con aproximaciones arbitrarias y_0 y $y_1 \in H$. Considerando y_k y y_{k-1} prefijados, eligiremos los parámetros α_{k+1} y τ_{k+1} de la condición de mínimo de la norma del error z_{k+1} en

H_D, es decir de la condición de optimización local por un paso según el esquema de tres capas.

Resolveremos este problema bajo la única suposición sobre la definición positiva del operador $DB^{-1}A$. Para esto pasaremos a la ecuación para el error equivalente $x_k = D^{1/2}z_k$:

$$x_{k+1} = \alpha_{k+1}(E - \tau_{k+1}C)x_k + (1 - \alpha_{k+1})x_{k-1},$$
$$C = D^{1/2}B^{-1}AD^{-1/2}. \qquad (6)$$

Para reducir los cálculos introduzcamos las notaciones

$$1 - \alpha_{k+1} = a, \quad \tau_{k+1}\alpha_{k+1} = b \qquad (7)$$

y volvamos a escribir (6) en la forma siguiente:

$$x_{k+1} = x_k - a(x_k - x_{k-1}) - bCx_k. \qquad (8)$$

El problema se plantea así: elegir a y b de la condición de mínimo de la norma de x_{k+1} en H. Calculemos la norma de x_{k+1}. De (8) obtenemos

$$\|x_{k+1}\|^2 = \|x_k\|^2 + a^2\|x_k - x_{k-1}\|^2 + b^2\|Cx_k\|^2 -$$
$$-2a(x_k, x_k - x_{k-1}) - 2b(Cx_k, x_k) + 2ab(Cx_k, x_k - x_{k-1})$$

Igualando a cero las derivadas parciales según a y b, obtendremos un sistema con respecto a los parámetros a y b

$$\|x_k - x_{k-1}\|^2 a + (Cx_k, x_k - x_{k-1})b = (x_k, x_k - x_{k-1}),$$
$$(Cx_k, x_k - x_{k-1})a + \|Cx_k\|^2 b = (Cx_k, x_k). \qquad (9)$$

El determinante del sistema es igual a $\|x_k - x_{k-1}\|^2 \times \|Cx_k\|^2 - (Cx_k, x_k - x_{k-1})^2$ y en virtud de la desigualdad de Cauchy—Buniakovsvi se anula solamente entonces, cuando $x_k - x_{k-1}$ es proporcionala Cx_k: $x_k - x_{k-1} = dCx_k$. En este caso las ecuaciones del sistema son proporcionales, y el mismo se reduce a la ecuación

$$(b + ad)\|Cx_k\|^2 = (Cx_k, x_k). \qquad (10)$$

Como con esto (8) tiene la forma $x_{k+1} = x_k - (b + ad)Cx_k$, entonces, poniendo $a = 0$ en (10), obtendremos de (7) y (10)

$$\alpha_{k+1} = 1, \quad \tau_{k+1} = \frac{(Cx_k, x_k)}{(Cx_k, Cx_k)}. \qquad (11)$$

Si el determinante no es igual a cero, entonces, resolviendo el sistema (9) obtendremos

$$a = \frac{\| Cx_k \|^2 (x_k, x_k - x_{k-1}) - (Cx_k, x_k)(Cx_k, x_k - x_{k-1})}{\| x_k - x_{k-1} \|^2 \| Cx_k \|^2 - (Cx_k, x_k - x_{k-1})^2},$$

$$b = \frac{(Cx_k, x_k)}{(Cx_k, Cx_k)}(1 - a) + \frac{(Cx_k, x_{k-1})}{(Cx_k, Cx_k)} a.$$

De aquí, utilizando las notaciones (7), hallaremos las fórmulas para los parámetros α_{k+1} y τ_{k+1}:

$$\alpha_{k+1} = \frac{(Cx_k, x_k - x_{k-1})(Cx_k, x_{k-1}) - (x_{k-1}, x_k - x_{k-1})(Cx_k, Cx_k)}{(Cx_k, Cx_k)(x_k - x_{k-1}, x_k - x_{k-1}) - (Cx_k, x_k - x_{k-1})^2}, \tag{12}$$

$$\tau_{k+1} = \frac{(Cx_k, x_k)}{(Cx_k, Cx_k)} + \frac{1 - \alpha_{k+1}}{\alpha_{k+1}} \frac{(Cx_k, x_{k-1})}{(Cx_k, Cx_k)}, \quad k = 1, 2, \ldots$$

Las fórmulas (11) obtenidas anteriormente pueden considerarse como un caso particular de las fórmulas generales (12), poniendo $\alpha_{k+1} = 1$, si el denominador en la expresión para α_{k+1} se anula.

Las fórmulas (12) son más complejas que las fórmulas para los parámetros α_{k+1} y τ_{k+1} del método de las direcciones conjugadas, obtenidas en el § 3. Aquí es necesario calcular productos escalares complementarios. Sin embargo, el proceso iterativo (5), (12) construido aquí es menos propenso a la influencia de los errores de redondeo, los errores admitidos en los pasos anteriores se extinguen.

La relación entre los métodos de tres capas localmente óptimos y los métodos de las direcciones conjugadas la establece el siguiente teorema.

TEOREMA 3. *Si para el método (5), (12) la aproximación inicial y_1 se elige de la manera siguiente:*

$$By_1 = (B - \tau_1 A) y_0 + \tau_1 f, \quad \tau_1 = \frac{(Dw_0, z_0)}{(Dw_0, w_0)}, \tag{13}$$

entonces en el caso del operador autoconjugado $DB^{-1}A$ este método coincide con el método de las direcciones conjugadas.

Realizaremos la demostración por inducción. De la condición del teorema se sigue que las aproximaciones y_1 obtenidas aquí y en el método de las direcciones conjugadas, coinciden. Supongamos que coinciden las aproximaciones $y_1, y_2, \ldots, y_k$. Demostremos que y_{k+1} construido según las fórmulas (5), (12), coincide con la aproximación y_{k+1} del método de las direcciones conjugadas.

De las suposiciones hechas se deduce que los parámetros iterativos $\tau_1, \tau_2, \ldots, \tau_k$ y $\alpha_2, \alpha_3, \ldots, \alpha_k$ también coinciden. Si mostramos que coinciden además los parámetros τ_{k+1} y α_{k+1} en estos métodos, entonces la afirmación del teorema 3 estará demostrada.

Puesto que $y_1, y_2, \ldots, y_k$ son las aproximaciones iterativas del método de las direcciones conjugadas, entonces en virtud del lema son cumplidas las condiciones

$$(Cx_j, x_i) = 0, \; j = 0, 1, \ldots, i-1, i = 1, 2, \ldots, k. \quad (14)$$

Sustituyendo (14) con $j = k-1$ e $i = k$ en (12) y utilizando la autoconjugación del operador C, obtendremos

$$\alpha_{k+1} = \frac{(x_{k-1}, x_k - x_{k-1})(Cx_k, Cx_k)}{(Cx_k, x_k)^2 - \| Cx_k \|^2 \| x_k - x_{k-1} \|^2}, \quad \tau_{k+1} = \frac{(Cx_k, x_k)}{(Cx_k, Cx_k)}. \quad (15)$$

Así, los parámetros τ_{k+1} del método localmente óptimo y del método de las direcciones conjugadas coinciden. Queda por mostrar que coinciden los parámetros α_{k+1}.

De (6) y (13) obtendremos

$$x_k - x_{k-1} = (\alpha_k - 1)(x_{k-1} - x_{k-2}) - \alpha_k \tau_k C x_{k-1}, \quad k = 2, 3, \ldots \quad (16)$$

$$x_1 - x_0 = -\tau_1 C x_0.$$

De (16) se desprende que la diferencia $x_k - x_{k-1}$ es una combinación lineal de $Cx_0, Cx_1, \ldots, Cx_{k-1}$ y tiene la forma siguiente:

$$x_k - x_{k-1} = -\alpha_k \tau_k C x_{k-1} + \sum_{j=0}^{k-2} \beta_j C x_j, \; k \geqslant 2,$$
$$x_1 - x_0 = -\tau_1 C x_0, \quad (17)$$

donde los coeficientes β_j se expresan mediante $\tau_1, \tau_2, \ldots, \ldots, \tau_{k-1}$ y $\alpha_2, \alpha_3, \ldots, \alpha_{k-1}$. Multiplicando escalarmente los miembros primero y segundo de (17) por x_{k-1} y $x_k - x_{k-1}$ y teniendo en cuenta (14), obtendremos

$$(x_{k-1}, x_k - x_{k-1}) = -\alpha_k \tau_k (Cx_{k-1}, x_{k-1}),$$
$$\| x_k - x_{k-1} \|^2 = \alpha_k \tau_k (Cx_{k-1}, x_{k-1}). \quad (18)$$

Sustituyendo (18) en la expresión (15) para α_{k+1} y teniendo en cuenta la fórmula para el parámetro τ_{k+1}, obtenemos

$$\alpha_{k+1} = \frac{\alpha_k \tau_k (Cx_{k-1}, x_{k-1})(Cx_k, Cx_k)}{\alpha_k \tau_k (Cx_{k-1}, x_{k-1})(Cx_k, Cx_k) - (Cx_k, x_k)^2} = \\ = \left(1 - \frac{\tau_{k+1}(Cx, x_k)}{\tau_k (Cx_{k-1}, x_{k-1})} \frac{1}{\alpha_k}\right)^{-1},$$

lo que coincide con la fórmula para el parámetro α_{k+1} en el método de las direcciones conjugadas. El teorema está demostrado.

Sustituyendo $x_k = D^{1/2} z_k$ y $C = D^{-1/2}(DB^{-1}A) D^{-1/2}$ en (12), obtenemos el siguiente tipo de fórmulas para los parámetros α_{k+1} y τ_{k+1}:

$$\alpha_{k+1} = \frac{(Dw_k, z_k - z_{k-1})(Dw_k, z_{k-1}) - (Dz_{k-1}, y_k - y_{k-1})(Dw_k, w_k)}{(Dw_k, w_k)(D(z_k - z_{k-1}), y_k - y_{k-1}) - (Dw_k, z_k - z_{k-1})^2}, \tag{19}$$

$$\tau_{k+1} = \frac{(Dw_k, z_k)}{(Dw_k, w_k)} + \frac{1 - \alpha_{k+1}}{\alpha_{k+1}} \frac{(Dw_k, z_{k-1})}{(Dw_k, w_k)}.$$

Si introducimos las notaciones para los productos escalares

$$a_k = (Dw_k, z_k), \quad b_k = (Dw_k, z_{k-1}), \quad c_k = \\ = (Dz_k, y_k - y_{k-1}),$$

$$d_k = (Dz_{k-1}, y_k - y_{k-1}), \quad e_k = (Dw_k, w_k),$$

entonces las fórmulas (19) se vuelven a escribir en la forma

$$\alpha_{k+1} = \frac{(a_k - b_k) b_k - d_k e_k}{(c_k - d_k) e_k - (a_k - b_k)^2}, \quad k = 1, 2, \ldots, \alpha_1 = 1,$$

$$\tau_{k+1} = \frac{a_k}{e_k} + \frac{1 - \alpha_{k+1}}{\alpha_{k+1}} \frac{b_k}{e_k}, \quad k = 0, 1, \ldots$$

Citemos las expresiones para a_k, b_k, c_k, d_k, y e_k para elecciones concretas del operador D:

1) $D = A$, $A = A^*$.
   $a_k = (w_k, r_k)$, $b_k = (w_k, r_{k-1})$, $c_k = (r_k, y_k - y_{k-1})$,
   $d_k = (r_{k-1}, y_k - y_{k-1})$, $e_k = (Aw_k, w_k)$.
2) $D = A^*A$
   $a_k = (Aw_k, r_k)$, $b_k = (Aw_k, r_{k-1})$,
   $c_k = (r_k, r_k - r_{k-1})$,
   $d_k = (r_{k-1}, r_k - r_{k-1})$, $e_k = (Aw_k, Aw_k)$.
3) $D = A^*B^{-1}A$, $B = B^*$.
   $a_k = (Aw_k, w_k)$, $b_k = (Aw_k, w_{k-1})$,
   $c_k = (w_k, r_k - r_{k-1})$,
   $d_k = (w_{k-1}, r_k - r_{k-1})$, $e_k = (B^{-1}Aw_k, Aw_k)$.

§ 5. Aceleración de la convergencia de los métodos de dos capas en el caso autoconjugado

1. **Algoritmo del proceso de aceleración.** En el punto 5 del § 1 fue establecido que en el caso del operador autoconjugado $DB^{-1}A$ los métodos gradientales de dos capas poseen la propiedad asintótica. La última se manifiesta en que para grandes números de iteraciones la velocidad de convergencia del método disminuye esencialmente en comparación con el comienzo de las iteraciones. Fue también mostrado que para grandes números de iteraciones los errores examinados después de una iteración, se hacen casi proporcionales.

Utilizando esta propiedad, construiremos ahora un proceso para acelerar la convergencia de los métodos gradientales de dos capas.

Para resolver la ecuación

$$Au = f \tag{1}$$

examinaremos el método iterativo gradiental de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f,\ k=0,\ 1,\ \ldots,\ y_0 \in H, \tag{2}$$

$$\tau_{k+1} = \frac{(Dw_k,\ z_k)}{(Dw_k,\ w_k)},\ k=0,\ 1,\ \ldots \tag{3}$$

Sea el operador $DB^{-1}A$ autoconjugado en H. Entonces el método iterativo posee la propiedad asintótica, y para un número de iteraciones k suficientemente grande tiene lugar la igualdad aproximada

$$z_{k+2} \approx \rho^2 z_k,\quad z_k = y_k - u. \tag{4}$$

Examinemos primeramente el caso, cuando en (4) se cumple la igualdad estricta, es decir $z_{k+2} = \rho^2 z_k$. Construyamos por las aproximaciones ya halladas y_k y y_{k+2} una nueva aproximación por medio de la fórmula

$$y = \alpha y_{k+2} + (1-\alpha)\, y_k,\quad \alpha = 1/(1-\rho^2). \tag{5}$$

Para el error $z = y - u$ obtendremos

$$z = \alpha z_{k+2} + (1-\alpha)\, z_k = (\alpha\rho^2 + 1 - \alpha)\, z_k =$$
$$= [1 - \alpha(1-\rho^2)]\, z_k = 0.$$

Por consiguiente, en caso de cumplirse estrictamente la igualdad (4) la combinación lineal (5) de las aproximaciones y_k y y_{k+2} da la solución exacta de la ecuación (1).

Como fue señalado en el § 1, el cumplimiento de la igualdad exacta en (4) es un caso excepcional, el cual tiene lugar solamente para una aproximación inicial especial. En el caso general tiene lugar la igualdad aproximada (4), y los razonamientos realizados más arriba permiten confiar en que alguna combinación lineal de y_k y y_{k+2} dará una buena aproximación a la solución del problema inicial.

Hallemos la mejor entre estas combinaciones lineales. Sean y_k, y_{k+1} y y_{k+2} aproximaciones iterativas, obtenidas por las fórmulas (2) y (3). Buscaremos una nueva aproximación y mediante la fórmula

$$y = \alpha y_{k+2} + (1 - \alpha)\, y_k. \tag{6}$$

Pongamos el problema de elegir el parámetro α de manera tal que la norma del error $z = y - u$ en H_D sea mínima.

Primeramente, utilizando el esquema (2), eliminamos y_{k+2} de (6). Obtendremos

$$By_{k+2} = (B - \tau_{k+2}A)\, y_{k+1} + \tau_{k+2}\, f$$

y después de la sustitución de y_{k+2} en (6) tendremos

$$By = \alpha\,(B - \tau_{k+2}A)\, y_{k+1} + (1 - \alpha)\, By_k + \alpha\tau_{k+2}\, f, \tag{7}$$

donde y_{k+1} se encuentra por el esquema de dos capas

$$By_{k+1} = (B - \tau_{k+1}A)\, y_k + \tau_{k+1}\, f. \tag{8}$$

Si consideramos que y_k es la aproximación inicial dada, entonces el esquema (7), (8) coincide con el esquema del método de las direcciones conjugadas, al mismo tiempo los parámetros τ_{k+1} y τ_{k+2} coinciden con los respectivos parámetros del método de las direcciones conjugadas. De la teoría de este método se deduce (punto 1, § 4, la fórmula (3)), que el valor óptimo del parámetro α se determina por la fórmula

$$\alpha = \frac{1}{1 - \dfrac{\tau_{k+2}}{\tau_{k+1}} \dfrac{(Dw_{k+1},\ z_{k+1})}{(Dw_k,\ z_k)}} \tag{9}$$

De esta forma, el problema planteado sobre la mejor elección del parámetro α está resuelto. Las fórmulas (6) y (9) determinan el procedimiento acelerador.

Notemos que para hallar y se puede utilizar no solamente la fórmula (6), sino calcular y por el siguiente esquema de dos capas:

$$B\bar{y}_{k+1} = (B - \bar{\tau}_{k+1}A)\, y_k + \bar{\tau}_{k+1}f,$$
$$By = (B - \bar{\tau}_{k+2}A)\, \bar{y}_{k+1} + \bar{\tau}_{k+2}f,$$

donde $\bar{\tau}_{k+1}$ y $\bar{\tau}_{k+2}$ son las raíces de la ecuación

$$\tau^2 - \alpha\,(\tau_{k+1} + \tau_{k+2})\,\tau + \alpha\tau_{k+1}\tau_{k+2} = 0.$$

En calidad de $\bar{\tau}_{k+1}$ se debe tomar la raíz mínima.

El empleo de (10) en lugar de (6) permite no aumentar el volumen de información intermedia a recordar.

2. Estimación de la efectividad. Examinemos ahora la efectividad del procedimiento de aceleración. Antes de calcular la norma del error $z = y - u$ en H_D, transformemos la expresión (9) para α.

El cambio $z_k = D^{-1/2}x_k$ en (9) da

$$\alpha = \left(1 - \frac{\tau_{k+2}}{\tau_{k+1}} \frac{(Cx_{k+1},\ x_{k+1})}{(Cx_k,\ x_k)}\right)^{-1},$$
$$C = D^{1/2}B^{-1}AD^{-1/2}. \qquad (11)$$

De (10) y (11) del § 1 tenemos

$$\|x_{k+1}\| = \rho_{k+1}\|x_k\|, \quad \rho^2_{k+1} = 1 - \frac{(Cx_k,\ x_k)^2}{(Cx_k,\ Cx_k)\,\|x_k\|^2} \qquad (12)$$

De la fórmula (9) del § 1 obtenemos

$$\tau_{k+1} = \frac{(Cx_k,\ x_k)}{(Cx_k,\ Cx_k)}. \qquad (13)$$

Utilizando (12) y (13) hallaremos

$$\frac{\tau_{k+2}}{\tau_{k+1}} \frac{(Cx_{k+1},\ x_{k+1})}{(Cx_k,\ x_k)} = \frac{1-\rho^2_{k+2}}{1-\rho^2_{k+1}}\,\rho^2_{k+1}.$$

Sustituyendo esta expresión en (11), obtenemos

$$\alpha = \frac{1-\rho^2_{k+1}}{1-2\rho^2_{k+1}+\rho^2_{k+1}\rho^2_{k+2}}. \qquad (14)$$

Calculemos ahora la norma del error $z = y - u$ en H_D. De (6) obtenemos

$$z = \alpha z_{k+2} + (1 - \alpha)\, z_k.$$

De aquí para el error equivalente $z_k = D^{1/2}x_k$ y $x = D^{1/2}z$ tendremos $x = \alpha x_{k+2} + (1 - \alpha)\, x_k$. Calculemos la norma de x en H. Obtenemos

$$\| x \|^2 = \alpha^2 \| x_{k+2} \|^2 + 2\alpha (1 - \alpha) (x_{k+2}, x_k) (1 - \alpha)^2 \times \\ \times \| x_k \|^2.$$

De la igualdad $(x_{k+2}, x_k) = \| x_{k+1} \|^2$ demostrada en el punto 5 del § 1 se deduce que

$$\| x \|^2 = \alpha^2 \| x_{k+2} \|^2 + 2\alpha (1 - \alpha) \| x_{k+1} \|^2 + \\ + (1 - \alpha)^2 \| x_k \|^2.$$

Sustituyendo aquí la expresión (14) para α y utilizando (12), obtendremos

$$\| x \|^2 = \frac{\rho_{k+2}^2 - \rho_{k+1}^2}{\rho_{k+1}^2 (1 - 2\rho_{k+1}^2 + \rho_{k+1}^2 \rho_{k+2}^2)} \| x_{k+2} \|^2 < \| x_{k+2} \|^2. \tag{15}$$

Ya que $\rho_{k+1} \leqslant \rho_{k+2} \leqslant \rho < 1$, entonces

$$1 - 2\rho_{k+1}^2 \rho_{k+1}^2 \rho_{k+2}^2 \geqslant (1 - \rho^2)^2,$$

por consiguiente, para la norma de x tiene lugar la estimación

$$\| x \|^2 \leqslant \left(\frac{\rho_{k+2}^2}{\rho_{k+1}^2} - 1 \right) \frac{\| x_{k+2} \|^2}{(1 - \rho^2)^2}.$$

En virtud de la propiedad asintótica para los números k suficientemente grandes tenemos $\rho_{k+1} \approx \rho_{k+2}$, por eso el efecto del procedimiento acelerador será significativo.

Señalemos que aunque una aceleración efectiva de la convergencia tiene lugar para grandes números de iteraciones k, este método también se puede utilizar para cualquier número de iteraciones. Se recomienda de vez en cuando interrumpir el proceso de iteraciones realizado según el esquema de dos capas (2), (3) y calcular la nueva aproximación por el procedimiento propuesto. El proceso de iteraciones puede terminarse en el cálculo de tal aproximación, si para la aproximación y_{k+2} hallada se cumple la desigualdad

$$\frac{\rho_{k+2}^2 - \rho_{k+1}^2}{\rho_{k+1}^2 (1 - 2\rho_{k+1}^2 + \rho_{k+1}^2 \rho_{k+2}^2)} \| z_{k+2} \|_D^2 \leqslant \varepsilon^2 \| z_0 \|_D^2.$$

En efecto, en este caso obtendremos en virtud de (15) $\| y - u \|_D \leqslant \varepsilon \| y_0 - u \|_D$ es decir será alcanzada la exactitud exigida ε.

3. Ejemplo. Para ilustrar la efectividad del procedimiento propuesto de aceleración de la convergencia de los métodos gradientales de dos capas examinaremos la solución del problema modelo por el método implícito del descenso más rápido. En calidad de ejemplo tomaremos el problema de Dirichlet de diferencias para la ecuación de Laplace en la red cuadrada $\bar{\omega} = \{x_{ij} = (ih, jh),\ 0 \leqslant i \leqslant N,\ 0 \leqslant j \leqslant N,\ h = 1/N\}$ en el cuadrado unidad

$$\Lambda u = \Lambda_1 u + \Lambda_2 u = 0,\ x \in \omega, \quad u|_\gamma = 0, \tag{16}$$
$$\Lambda_\alpha u = u_{\bar{x}_\alpha x_\alpha}, \quad \alpha = 1,\ 2.$$

Introduzcamos el espacio H que consiste de las funciones reticulares definidas sobre ω, con el producto escalar

$$(u,\ v) = \sum_{x \in \omega} u(x)\, v(x)\, h^2.$$

Definiremos el operador A de la forma siguiente: $A = A_1 + A_2$, $A_\alpha y = -\Lambda_\alpha v$, $y \in H$, donde $v(x) = y(x)$ para $x \in \omega$ y $v|_\gamma = 0$.

Escribamos el problema (16) en forma de la ecuación operacional

$$Au = f, \quad f = 0. \tag{17}$$

En calidad de operador B eligiremos el siguiente operador factorizado: $B = (E + \omega A_1)(E + \omega A_2)$, $\omega > 0$, donde ω es un parámetro iterativo.

Puesto que los operadores A_1 y A_2 son autoconjugados y conmutan en H, entonces los operadores A y B son autoconjugados en H. Además, es fácil mostrar que los operadores A y B son definidos positivos en H. Por consiguiente, para resolver la ecuación (17) se puede utilizar el método implícito del descenso más rápido

$$B\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad \tau_{k+1} = \frac{(w_k,\ r_k)}{(Aw_k,\ w_k)},$$
$$k = 0,\ 1,\ \ldots \tag{18}$$

En este método $D = A$ y $DB^{-1}A = AB^{-1}A$. Como el operador $DB^{-1}A$ es autoconjugado en H, entonces para el método a examinar tiene lugar la propiedad asintótica. De la teoría del método del descenso más rápido (el punto 1 del § 2)

se sigue que la velocidad de convergencia del método en el caso dado está determinada por la relación $\xi = \gamma_1/\gamma_2$, donde γ_1 y γ_2 son las constantes de equivalencia energética de los operadores A y B: $\gamma_1 B \leqslant A \leqslant \gamma_2 B$, $\gamma_1 > 0$.

Por lo tanto el parámetro iterativo ω se elige de la condición de maximalidad de ξ. En el § 2 del cap. XI será mostrado que el valor óptimo ω se determina por la fórmula

$$\omega = \frac{1}{\sqrt{\delta\Delta}}, \quad \delta = \frac{4}{h^2}\operatorname{sen}^2\frac{\pi h}{2}, \quad \Delta = \frac{4}{h^2}\cos^2\frac{\pi h}{2},$$

con esto

$$\gamma_1 = \frac{2\delta}{(1+\sqrt{\eta})^2}, \quad \gamma_2 = \frac{(\Delta+\delta)\sqrt{\eta}}{(1+\sqrt{\eta})^2}, \quad \xi = \frac{2\sqrt{\eta}}{1+\eta}, \quad \eta = \frac{\delta}{\Delta}.$$

Para el ejemplo a examinar

$$\omega = \frac{h^2}{2\operatorname{sen}\pi h}, \quad \gamma_1 = \frac{2}{h^2}\frac{\operatorname{sen}^2\pi h}{1+\operatorname{sen}\pi h}, \quad \gamma_2 = \frac{2}{h^2}\frac{\operatorname{sen}\pi h}{1+\operatorname{sen}\pi h},$$

$$\xi = \operatorname{sen}\pi h.$$

Citemos los resultados de los cálculos, cuando la aproximación inicial y_0 se eligió igual a $y_0(x) = e^{(x_1-x_2)}$ para $x \in \omega$, $y_0|_\gamma = 0$. La exactitud exigida ε se tomó igual a 10^{-4}, $N = 40$.

En la tabla 8 para algunos números de iteraciones k se dan: $\|z_k\|_D/\|z_0\|_D$ — exactitud relativa de la k-ésima ite-

Tabla 8

k	$\|z_k\|_D/\|z_0\|_D$	ρ_k	$\gamma_1^{(k)}$	$\gamma_2^{(k)}$	τ_k
1	$3{,}6\cdot10^{-1}$	0,36203	—	—	$5{,}392\cdot10^{-3}$
2	$2{,}3\cdot10^{-1}$	0,63810	77,31858	236,1883	$7{,}809\cdot10^{-3}$
3	$1{,}8\cdot10^{-1}$	0,76998	40,59796	232,1435	$6{,}911\cdot10^{-3}$
4	$1{,}4\cdot10^{-1}$	0,81178	26,87824	233,4976	$8{,}644\cdot10^{-3}$
. . . . .					
26	$3{,}9\cdot10^{-3}$	0,85175	18,27141	230,5962	$8{,}876\cdot10^{-3}$
27	$3{,}4\cdot10^{-3}$	0,85178	18,26983	230,6607	$7{,}338\cdot10^{-3}$
28	$2{,}9\cdot10^{-3}$	0,85183	18,27026	230,7191	$8{,}872\cdot10^{-3}$
29	$2{,}4\cdot10^{-3}$	0,85186	18,26895	230,7771	$7{,}335\cdot10^{-3}$
. . . . .					
46	$1{,}6\cdot10^{-4}$	0,85226	18,26677	231,4121	$8{,}845\cdot10^{-3}$
47	$1{,}4\cdot10^{-4}$	0,85227	18,26632	231,4375	$7{,}318\cdot10^{-3}$
48	$1{,}2\cdot10^{-4}$	0,85229	18,26664	231,4612	$8{,}843\cdot10^{-3}$
49	$9{,}9\cdot10^{-5}$	0,85230	18,26623	231,4849	$7{,}317\cdot10^{-3}$

ración, $\rho_k = \| z_k \|_D / \| z_{k-1} \|_D$ es la magnitud que caracteriza la disminución de la norma del error al pasar de la $(k-1)$-ésima iteración a la k-ésima, $\gamma_1^{(k)}$ y $\gamma_2^{(k)}$ son las aproximaciones para γ_1 y γ_2, las cuales se encuentran como las raíces de la ecuación cuadrática

$$(1 - \tau_k \gamma)(1 - \tau_{k-1}\gamma) = \rho_k \rho_{k-1}, \quad k = 2, 3, \ldots,$$

y los parámetros iterativos τ_k.

La exactitud prefijada ε fue alcanzada después de ser cumplidas 49 iteraciones según el esquema (18). Para $\varepsilon = 10^{-4}$ el número teórico de iteraciones es igual a 59. Los valores citados en la tabla para ρ_k ilustran bien la propiedad asintótica del método. Se ve, que con el crecimiento del número de iteraciones ocurre una disminución de la velocidad de convergencia del método. La exactitud $4 \cdot 10^{-3}$ fue alcanzada tras 26 iteraciones y en el aumento de la exactitud aún en 40 veces se exigió realizar complementariamente 23 iteraciones. La magnitud ρ_k crece monótonamente y para $k = 26$ tenemos $\rho_{k+1} - \rho_k \approx 3 \cdot 10^{-5}$. Los parámetros iterativos τ_k y τ_{k+2} se hacen casi iguales.

Para comparar los valores aproximados $\gamma_1^{(k)}$ y $\gamma_2^{(k)}$ con los exactos citemos a γ_1 y γ_2:

$$\gamma_1 = 18{,}26556, \quad \gamma_2 = 232{,}8036.$$

Después de efectuar las 49 iteraciones γ_1 fue hallado con una exactitud del 0,004%, y γ_2 con una exactitud del 0,6%.

Para acelerar la convergencia del método fue aplicado el procedimiento de aceleración descrito en el punto 1. Por las aproximaciones y_{26} y y_{28}, halladas según el esquema (18), mediante las fórmulas (6) y (9) fue construida una nueva aproximación y. La exactitud dada $\varepsilon = 10^{-4}$ fue alcanzada. La aplicación del método de aceleración de la convergencia de los métodos gradientales de dos capas construido en este párrafo, permitió disminuir el número de iteraciones exigidas para el ejemplo examinado en 1,8 veces aproximadamente.

Capítulo IX

Métodos iterativos triangulares

En este capítulo se estudian los métodos iterativos implícitos de dos capas, en los cuales al operador B le corresponden matrices triangulares. En el § 1 se examina el método de Seidel, para el cual se formulan las condiciones suficientes de convergencia. En el § 2 se investiga el método de relajación superior. Aquí se da la elección del parámetro iterativo y se obtiene la estimación para el radio espectral del operador de transición. En el § 3 es examinado el esquema iterativo general de los métodos triangulares, está indicada la elección del parámetro iterativo y demostrada la convergencia del método en la norma de H_A.

§ 1. Método de Seidel

1. **Esquema iterativo del método.** En los capítulos anteriores fue expuesta la teoría de los métodos iterativos de dos y tres capas, aplicados para encontrar una solución aproximada de la ecuación operacional lineal de primer género

$$Au = f. \tag{1}$$

Esta teoría indica la elección de los parámetros iterativos y da la estimación para el número de iteraciones de los correspondientes métodos, al mismo tiempo la teoría utiliza el mínimo de información de carácter general con respecto a los operadores del esquema iterativo. Al renunciar el estudio de la estructura concreta de los operadores del esquema iterativo permite desarrollar la teoría desde un único punto de vista y construir métodos iterativos implícitos, optimales en la clase de los operadores B.

En la teoría general de los métodos iterativos fue mostrado que la efectividad de un método depende de manera esen-

cial de la elección del operador B. De la elección del operador B dependen tanto el número de iteraciones que es necesario cumplir para alcanzar una exactitud dada ε, como el número de operaciones aritméticas, requeridas para realizar un paso iterativo. Cada una de estas magnitudes por separado no puede servir como criterio de la efectividad de un método iterativo. Aclaremos esta afirmación. Sean los operadores A y B autoconjugados y definidos positivos en H. De la teoría de los métodos iterativos se deduce que si en calidad de operador D se toma uno de los operadores A, B o $AB^{-1}A$, entonces el número de iteraciones para los métodos iterativos examinados en los cap. VI-VIII (de Chebishev, de iteración simple, de los métodos de tipo variacional, etc.) se determina por el cociente $\xi = \gamma_1/\gamma_2$, donde γ_1 y γ_2 son las constantes de equivalencia energética de los operadores A y B:

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B.$$

Por eso, si elegimos $B = A$, entonces obtendremos el valor máximo posible $\xi = 1$, y los métodos iterativos darán la solución exacta de la ecuación (1) por una iteración para cualquier aproximación inicial. Por consiguiente, la elección indicada del operador B minimiza el número de iteraciones. Sin embargo para realizar este único paso iterativo se exige invertir el operador B, es decir el operador A. Es evidente, que para esto el número de operaciones aritméticas será máximo.

Por otra parte, para los esquemas explícitos con $B = E$ se exige un número mínimo de operaciones aritméticas en una iteración, pero el número de iteraciones se hace demasiado grande.

De esta forma, aparece el problema de la elección óptima del operador B de la condición de minimizar el volumen total de trabajo computacional, el cual debe ser cumplido para obtener la solución con una exactitud dada.

Es natural, que en tal planteamiento general este problema no puede ser resuelto. En la actualidad el desarrollo de los métodos iterativos va por el camino de la construcción de operadores B fácilmente inversibles, entre los cuales se eligen los operadores con la mejor relación γ_1/γ_2. A los operadores fácilmente inversibles o económicos usualmente se relacionan aquellos operadores, cuya inversión se realiza en un número de operaciones aritméticas, proporcional o casi proporcional al número de incógnitas. Ejemplos de tales operadores son los operadores, a los cuales les corresponden

matrices diagonales, tridiagonales y triangulares, y también sus productos. En calidad de un ejemplo más complejo citemos el operador de Laplace de diferencias en rectángulo, el cual, como fue mostrado en el capítulo IV, se puede invertir con métodos directos con un pequeño gasto de operaciones aritméticas.

Se debe señalar que la utilización en calidad de operador B de operadores diagonales permite disminuir el número de iteraciones en comparación con el caso de un esquema iterativo explícito. Sin embargo, el orden asintótico de la dependencia del número de iteraciones del número de incógnitas del problema se mantiene igual que para un esquema explícito. Una dirección más perspectiva es la utilización de operadores B triangulares.

En este capítulo y en el capítulo X serán examinados los métodos iterativos implícitos de dos capas universales, en los cuales al operador B le corresponden matrices triangulares (métodos triangulares) o productos de matrices triangulares (método alternativo-triangular).

Comenzaremos el examen de estos métodos por el más simple—*el método de Seidel.*

Examinemos el sistema de ecuaciones algebraicas lineales (1) o en forma detallada

$$\sum_{j=1}^{M} a_{ij}u_j = f_i, \quad i = 1, 2, \ldots, M.$$

En este caso nosotros nos ocupamos de la ecuación (1) prefijada en el espacio de dimensión finita $H = H_M$.

Bajo la suposición de que los elementos diagonales de la matriz $A = (a_{ij})$ son diferentes de cero ($a_{ii} \neq 0$), el método iterativo de Seidel se escribe en la forma siguiente:

$$\sum_{j=1}^{i} a_{ij}y_j^{(k+1)} + \sum_{j=i+1}^{M} a_{ij}y_j^k = f_i.$$

$$i = 1, 2, \ldots, M, \qquad (2)$$

donde $y_j^{(k)}$ es la j-ésima componente de la aproximación iterativa de número k. En calidad de aproximación inicial se elige un vector arbitrario.

Comenzamos la determinación de la $(k+1)$-ésima iteración desde $i = 1$:

$$a_{11}y_1^{(k+1)} = -\sum_{j=2}^{M} a_{1j}y_j^{(k)} + f_1.$$

Ya que $a_{11} \neq 0$, entonces de aquí hallaremos $y_1^{(k+1)}$. Para $i = 2$ obtendremos

$$a_{22}y_2^{(k+1)} = -a_{21}y_1^{(k+1)} - \sum_{j=3}^{M} a_{2j}y_j^{(k)} + f_2.$$

Sean ya hallados $y_1^{(k+1)}$, $y_2^{(k+1)}$, ..., $y_{i-1}^{(k+1)}$. Entonces $y_i^{(k+1)}$ se encuentra de la ecuación

$$a_{ii}y_i^{(k+1)} = -\sum_{j=1}^{i-1} a_{ij}y_j^{(k+1)} - \sum_{j=i+1}^{M} a_{ij}y_j^{(k)} + f_i. \qquad (3)$$

De la fórmula (3) se ve, que el algoritmo del método de Seidel es extraordinariamente simple. El valor $y_i^{(k+1)}$ hallado por la fórmula (3) se sitúa en el lugar de $y_i^{(k)}$.

Estimemos el número de operaciones aritméticas que se exige para la realización de un paso iterativo. Si todos los a_{ij} no son iguales a cero, entonces los cálculos según la fórmula (3) exigen $M - 1$ operaciones de adición, $M - 1$ operaciones de multiplicación y una división. Por eso la realización de un paso iterativo se lleva a efecto en $2M^2 - M$ operaciones aritméticas.

Si en cada fila de la matriz A son diferentes de cero únicamente m elementos, y precisamente esta situación tiene lugar para las ecuaciones elípticas reticulares, entonces en la realización de un paso iterativo se exigen $2mM - M$ operaciones, o sea, el número de operaciones es proporcional al número de incógnitas M.

Escribamos ahora el método iterativo de Seidel (2) en forma matricial. Para eso representemos la matriz A en forma de la suma de una matriz diagonal, una triangular inferiormente y una triangular superiormente

$$A = \mathscr{D} + L + U, \qquad (4)$$

donde

$$L = \begin{Vmatrix} 0 & & & & \\ a_{21} & 0 & & & 0 \\ a_{31} & a_{32} & 0 & & \\ \cdot & \cdot & \cdot & & \\ \cdot & & \cdot & & \cdot \\ a_{M1} & a_{M2} & \dots & a_{M\,M-1} & 0 \end{Vmatrix},$$

$$U = \begin{Vmatrix} 0 & a_{12} & a_{13} & \dots & a_{1M} \\ & 0 & a_{23} & \dots & a_{2M} \\ & & \cdot & & \cdot \\ & & \cdot & \cdot & \cdot \\ & & & 0 & a_{M-1M} \\ 0 & & & & 0 \end{Vmatrix}$$

$$\mathscr{D} = \begin{Vmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \cdot & \\ & & & \cdot \\ & 0 & & a_{MM} \end{Vmatrix}$$

Designemos mediante $y_k = (y_1^{(k)}, y_2^{(k)}, \dots, y_M^{(k)})$ el vector de la k-ésima aproximación iterativa.

Utilizando estas notaciones, escribamos el método de Seidel en la forma

$$(\mathscr{D} + L)\, y_{k+1} + Uy_k = f, \quad k = 0, 1, \dots$$

Redujamos este esquema iterativo a la forma canónica de los esquemas de dos capas

$$(\mathscr{D} + L)\,(y_{k+1} - y_k) + Ay_k = f, \quad k = 0, 1, \dots$$
$$\dots, y_0 \in H. \quad (5)$$

Comparando (5) con la forma canónica

$$B\,\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f,\ k = 0, 1, \dots, y_0 \in H.$$

encontramos, que $B = \mathscr{D} + L$, $\tau_k \equiv 1$. El esquema (5) es implícito, el operador B es una matriz triangular y, por lo tanto, es autoconjugado en H.

Hemos examinado el así llamado método de Seidel puntual o escalar, suponiendo, que los elementos a_{ij} de la matriz A son números. Análogamente se construye el método de Seidel por bloques o vectorial para el caso, cuando a_{ii} es una matriz cuadrada, en general, de diferente dimensión, y a_{ij} para $i \neq j$ son las matrices rectangulares. En este caso y_i y f_i son los vectores, cuya dimensión corresponde a la dimensión de la matriz a_{ii}.

En la suposición de no degeneración de las matrices a_{ii} el método de bloques de Seidel se escribe en la forma (2) o en la forma canónica (5).

2. Ejemplos de aplicación del método. Examinemos la aplicación del método de Seidel para encontrar una solución aproximada del problema de Dirichlet de diferencias para la ecuación de Poisson y para una ecuación elíptica con coeficientes variables en un rectángulo.

Supongamos que sobre la red rectangular $\bar{\omega} = \{x_{ij} = (ih_1,\ ih_2) \in \bar{G},\ \ 0 \leqslant i \leqslant N_1,\ \ 0 \leqslant j \leqslant N_2,\ \ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1,\ 2\}$, introducida en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1,\ 2\}$, se exige hallar la solución del problema de Dirichlet de diferencias para la ecuación de Poisson

$$\Lambda y = \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha x_\alpha} = -\varphi(x),\ x \in \omega,\ y(k) = g(x),\ x \in \gamma, \qquad (6)$$

donde

$$\gamma = \{x_{ij} \in \Gamma\} \text{ es la frontera de la red } \bar{\omega}.$$

En este ejemplo las incógnitas son $y(i,\ j) = y(x_{ij})$ en los nodos interiores de la red. Si ordenamos las incógnitas de manera natural por las filas de la red ω, comenzando por la fila inferior, entonces el esquema en diferencias (6) puede ser escrito en la forma del siguiente sistema de ecuaciones algebraicas:

$$-\frac{1}{h_1^2} y(i-1,\ j) - \frac{1}{h_2^2} y(i,\ j-1) + \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y(i,\ j) - \\ - \frac{1}{h_1^2} y(i+1,\ j) - \frac{1}{h_2^2} y(i,\ j+1) = \varphi(i,\ j)$$

para $i = 1,\ 2,\ \ldots,\ N_1 - 1,\ j = 1,\ 2,\ \ldots,\ N_2 - 1$ y $y(x) = g(x)$ para $x \in \gamma$. Con esto las incógnitas $y(i-1,\ j)$ y $y(i,\ j-1)$ preceden a $y(i,\ j)$, mientras que $y(i+1,\ j)$ y $y(i,\ j+1)$ siguen después de $y(i,\ j)$. Ya que en cada ecuación están relacionadas no más de cinco incógnitas, entonces en cada fila de la matriz A son diferentes de cero no más de cinco elementos.

Para el sistema examinado el método puntual de Seidel tendrá la forma siguiente:

$$\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y_{k+1}(i,\ j) = \frac{1}{h_1^2} y_{k+1}(i-1,\ j) + \frac{1}{h_2^2} y_{k+1} \times \\ \times (i,\ j-1) + \frac{1}{h_1^2} y_k(i+1,\ j) + \frac{1}{h_2^2} y_k(i,\ j+1) + \varphi(i,\ j), \\ 1 \leqslant i \leqslant N_1 - 1,\ \ 1 \leqslant j \leqslant N_2 - 1,$$

al mismo tiempo $y_k(x) = g(x)$, $x \in \gamma$ para cualquier $k \geqslant 0$.

Los cálculos se comienzan por el punto $i = 1$, $j = 1$ y se continúan o bien por las filas, o por las columnas de la red ω. Para encontrar $y_{k+1}(i, j)$ se exigen 7 operaciones aritméticas y en total para realizar un paso iterativo se exigen $7M$ operaciones, donde $M = (N_1 - 1)(N_2 - 1)$ es el número de incógnitas en el problema.

Para el ejemplo examinado el operador B en el espacio de dimensión finita H de las funciones reticulares, definidas sobre ω, con el producto escalar $(u, v) = \sum_{x \in \omega} u(x) v(x) h_1 h_2$, $u, v \in H$ se define de la siguiente forma:

$$By = (\mathscr{D} + L) y = \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) \overset{\circ}{y}(i, j) - \frac{1}{h_1^2} \overset{\circ}{y}(i-1, j) -$$

$$- \frac{1}{h_2^2} \overset{\circ}{y}(i, j-1) = \left(\frac{1}{h_1^2} + \frac{1}{h_2^2}\right) \overset{\circ}{y} + \sum_{\alpha=1}^{2} \frac{1}{h_\alpha} \overset{\bullet}{y}_{\bar{x}_\alpha},$$

donde $y \in H$, $\overset{\circ}{y} \in \overset{\circ}{H}$ y $\overset{\circ}{y}(x) = y(x)$ para $x \in \omega$. Aquí $\overset{\circ}{H}$ es el conjunto de las funciones reticulares definidas sobre ω y que se anulan sobre γ.

Examinemos ahora *el método de bloques de Seidel.* Si designamos mediante $Y_j = (y(1, j), y(2, j), \ldots, y(N_1 - 1, j))$ el vector que consiste de las incógnitas en la fila j de la red, entonces, como fue mostrado en el § 1 del cap. I, el problema de diferencias (6) puede ser escrito en la forma del sistema tripuntual de ecuaciones vectoriales:

$$-Y_{j-1} + CY_j - Y_{j+1} = F_j, \quad j = 1, 2, \ldots, N_2 - 1, \quad (7)$$
$$Y_0 = F_0, \quad Y_{N_2} = F_{N_2},$$

donde C es una matriz cuadrada tridiagonal de dimensión $(N_1 - 1)(N_1 - 1)$ definida de la siguiente forma:

$$(CY_j)_i = (2y - h_2^2 y_{\bar{x}_1 x_1})_{ij}, \quad y_{0j} = y_{N_1 j} = 0.$$

Los miembros derechos F_j se determinan por las fórmulas

$$F_j = \left(h_2^2 \varphi(1, j) + \frac{h_2^2}{h_1^2} g(0, j),\ h_2^2 \varphi(2, j),\ \ldots\right.$$

$$\left. \ldots,\ h_2^2 \varphi(N_1 - 2, j),\ h_2^2 \varphi(N_1 - 1, j) + \frac{h_2^2}{h_1^2} g(N_1, j)\right)$$

para $j = 1, 2, \ldots, N_2 - 1$.

$F_j = (g(1, j), g(2, j), \ldots, g(N_1 - 1), j))$ para $j = 0, N_2$.

El método de bloques de Seidel para el sistema (7) tiene la forma

$$CY_j^{(k+1)} = Y_{j-1}^{(k+1)} + Y_{j+1}^{(k)} + F_j, \quad j = 1, 2, \ldots, N_2 - 1,$$
$$Y_0^{(k+1)} = F_0, \quad Y_{N_2}^{(k)} = F_{N_2}, \quad k = 0, 1, \ldots, \tag{8}$$

y para encontrar $Y_j^{(k+1)}$ se exige invertir la matriz tridiagonal C.

Si anotamos el esquema (8) por los puntos de la red, entonces obtendremos las fórmulas siguientes:

$$-\frac{1}{h_1^2} y_{k+1}(i-1, j) + \left(\frac{2}{h_1} + \frac{2}{h_2^2}\right) y_{k+1}(i, j) - \frac{1}{h_1^2}(i+1, j) =$$
$$= \frac{1}{h_2^2} y_{k+1}(i, j-1) + \frac{1}{h_2^2} y_k(i, j+1) + \varphi(i, j), \tag{9}$$
$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1,$$

al mismo tiempo $y_k(x) = g(x)$, $x \in \gamma$ para cualquier $k \geqslant 0$. Para encontrar y_{k+1} en la j-ésima fila es necesario resolver el problema de contorno tripuntual (9) del segundo miembro conocido, por ejemplo por el método de factorización, y situar la solución obtenida en el lugar de y_k en la j-ésima fila.

Al método de bloques de Seidel le corresponde el siguiente operador B:

$$By = \frac{1}{h_2^2} \overset{\circ}{y} + \frac{1}{h_2} \overset{\circ}{y}_{\bar{x}_2} + y_{\bar{x}_1 x_1}, \quad y \in H, \quad \overset{\circ}{y} \in \overset{\circ}{H}.$$

Supongamos ahora que sobre la red $\bar{\omega}$ hay que hallar la solución del problema de Dirichlet de diferencias para la ecuación elíptica con coeficientes variables

$$\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha(x) y_{\bar{x}_\alpha})_{x_\alpha} - d(x) y = -\varphi(x), \quad x \in \omega, \quad y(x) =$$
$$= g(x), \quad x \in \gamma, \tag{10}$$
$$0 \leqslant c_1 \leqslant a_\alpha(x) \leqslant c_2, \quad x \in \bar{\omega}, \quad \alpha = 1, 2, \quad 0 \leqslant d_1 \leqslant d(x) \leqslant$$
$$\leqslant d_2, \quad x \in \omega.$$

Para el problema a examinar bajo el ordenamiento de las incógnitas por las filas de la red el método puntual de Seidel tendrá la forma siguiente:

$$\left(\frac{a_1(i+1,\ j)+a_1(i,\ j)}{h_1^2}+\frac{a_2(i,\ j+1)+a_2(i,\ j)}{h_2^2}+\right.$$
$$\left.+d(i,\ j)\right)y_{k+1}(i,\ j)=\frac{a_1(i,\ j)}{h_1^2}y_{k+1}(i-1,\ j)+$$
$$+\frac{a_2(i,\ j)}{h_2^2}y_{k+1}(i,\ j-1)+\frac{a_1(i+1,\ j)}{h_1^2}y_k(i+1,\ j)+$$
$$+\frac{a_2(i,\ j+1)}{h_2^2}y_k(i,\ j+1)+\varphi_{ij}$$

para $i=1,\ 2,\ \ldots,\ N_1-1$ y $j=1,\ 2,\ \ldots,\ N_2-1$, al mismo tiempo $y_k(x)=g(x)$ con $x\in\gamma$ para cualquier $k\geqslant 0$.

El operador B en la forma canónica del esquema iterativo para el ejemplo dado se define de la siguiente forma:

$$By(x_{ij})=\left(\frac{a_1(i+1,\ j)}{h_1^2}+\frac{a_2(i,\ j+1)}{h_2^2}+d(i,\ j)\right)\overset{\circ}{y}(i,\ j)+$$
$$+\frac{a_1(i,\ j)}{h_1}\overset{\circ}{y}_{\bar{x}_1}+\frac{a_2(i,\ j)}{h_2}\overset{\circ}{y}_{\bar{x}_2},\quad y\in H,\quad \overset{\circ}{y}\in\overset{\circ}{H},$$

donde el espacio H y el conjunto $\overset{\circ}{H}$ están definidos más arriba.

3. Condiciones de convergencia suficientes. Formulemos ahora algunas condiciones de convergencia suficientes del método de Seidel. Nosotros necesitaremos el siguiente

TEOREMA 1. *Sea el operador A en la ecuación (1) autoconjugado y definido positivo en H. Entonces el proceso iterativo de dos capas*

$$B\frac{y_{k+1}-y_k}{\tau}+Ay_k=f,\quad k=0,\ 1,\ \ldots,\quad y_0\in H,\quad \tau>0, \tag{11}$$

converge en H_A, si el operador $B-0{,}5\tau A$ es definido positivo en H, es decir, está cumplida la condición

$$B>\frac{\tau}{2}A. \tag{12}$$

En realidad, de (11) obtenemos para el error $z_k=y_k-u$ el siguiente problema:

$$B\frac{z_{k+1}-z_k}{\tau}+Az_k=0,\quad k=0,\ 1,\ \ldots,\quad z_0=y_0-u. \tag{13}$$

Establezcamos para z_k una identidad energética fundamental. Pongamos en (13) z_k en la forma $z_k = \frac{1}{2}(z_{k+1} + z_k) - \frac{\tau}{2}\left(\frac{z_{k+1}-z_k}{\tau}\right)$ y obtendremos

$$\left(B - \frac{\tau}{2}A\right)\frac{z_{k+1}-z_k}{\tau} + \frac{1}{2}A(z_{k+1}+z_k) = 0.$$

Multipliquemos los miembros primero y segundo de esta igualdad escalarmente por $2(z_{k+1} - z_k)$ y tengamos en cuenta que el operador autoconjugado A tiene lugar la igualdad $(A(z_{k+1} + z_k), z_{k+1} - z_k) = (Az_{k+1}, z_{k+1}) - (Az_k, z_k)$. Como resultado obtendremos la identidad energética fundamental

$$2\tau\left(\left(B - \frac{\tau}{2}A\right)\frac{z_{k+1}-z_k}{\tau}, \frac{z_{k+1}-z_k}{\tau}\right) + \\ + \| z_{k+1} \|_A^2 - \| z_k \|_A^2 = 0.$$

De aquí y de las desigualdades $B - 0{,}5\tau A > 0$, $\tau > 0$ se infiere, que $\| z_{k+1} \|_A^2 \leqslant \| z_k \|_A^2$, es decir la sucesión $\{\| z_k \|_A^2\}$ no crece, está acotada inferiormente por cero y es convergente. Entonces de la identidad energética se sigue que

$$\lim_{k\to\infty}\left(\left(B - \frac{\tau}{2}A\right)\frac{z_{k+1}-z_k}{\tau}, \frac{z_{k+1}-z_k}{\tau}\right) = 0. \qquad (14)$$

Después, de la desigualdad $B - 0{,}5\tau A > 0$ se infiere que $\| z_{k+1} - z_k \| \to 0$ cuando $k \to \infty$. Observando de (13) que $A^{\frac{1}{2}} z_k = -A^{\frac{1}{2}} B(z_{k+1} - z_k)/\tau$, obtendremos $\| z_k \|_A^2 \leqslant \| A^{-1} \| \| B \|^2 \| z_{k+1} - z_k \|^2/\tau^2 \to 0$ para $k \to \infty$.

Formulemos una condición suficiente de convergencia del método de Seidel.

TEOREMA 2. *Si el operador A es autoconjugado y definido positivo en H, entonces el método de Seidel (4), (5) converge en H_A.*

En efecto, de (5) y del teorema 1 se deduce que es suficiente comprobar la desigualdad $\mathscr{D} + L - 0{,}5A > 0$. Ya que $A = A^*$, entonces en (4) tenemos $U = L^*$ y

$$((\mathscr{D} + L - 0{,}5A)x, x) = 0{,}5((\mathscr{D} + L - U)x, x) = \\ = 0{,}5(\mathscr{D}x, x).$$

Como A es un operador definido positivo, entonces para el método de Seidel puntual tenemos $a_{ii} > 0$, $1 \leqslant i \leqslant M$, y para el método de bloques de Seidel las matrices $a_{ii} = a_{ii}^* > 0$. Por consiguiente, $\mathscr{D} = \mathscr{D}^* > 0$. De esta forma $\mathscr{D} + L - 0{,}5A > 0$.

Citemos sin demostración otra condición de convergencia del método de Seidel.

TEOREMA 3. *Si el operador A es autoconjugado y no degenerado, y todos los $a_{ii} > 0$, entonces el método de Seidel converge para toda aproximación inicial cuando y sólo cuando A es un operador definido positivo.*

Para estimar la velocidad de convergencia del método de Seidel se utilizan suposiciones de distinto género.

Por ejemplo, si se cumple la condición

$$\sum_{j \neq i} |a_{ij}| \leqslant q |a_{ii}|, \quad i = 1, 2, \ldots, M, \quad q < 1, \tag{15}$$

entonces el método de Seidel converge con la velocidad de una progresión geométrica con denominador q, y para el error z_n tiene lugar la estimación $\| z_n \| \leqslant q^n \| z_0 \|$, donde

$$\| z_n \| = \max_{1 \leqslant i \leqslant M} |y_i^{(n)} - u_i|.$$

En efecto, de (3) obtenemos para el error $z_i^{(k)} = y_i^{(k)} - u_i$ la ecuación homogénea

$$a_{ii} z_i^{(k+1)} = -\sum_{j=1}^{i-1} a_{ij} z_j^{(k+1)} - \sum_{j=i+1}^{M} a_{ij} z_j^{(k)}.$$

De aquí hallaremos

$$|z_i^{(k+1)}| \leqslant \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| |z_j^{(k+1)}| + \sum_{j=i+1}^{M} \left| \frac{a_{ij}}{a_{ii}} \right| |z_j^{(k)}| \leqslant$$

$$\leqslant \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \| z_{k+1} \| + \sum_{j=i+1}^{M} \left| \frac{a_{ij}}{a_{ii}} \right| \| z_k \|. \tag{16}$$

De (15) obtenemos

$$\sum_{j=i+1}^{M} \left| \frac{a_{ij}}{a_{ii}} \right| \leqslant q - \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \leqslant q \left(1 - \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \right).$$

Sustituyendo esta estimación en (16), obtendremos la siguiente desigualdad:

$$|z_i^{(k+1)}| \leqslant \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \| z_{k+1} \| + q \left(1 - \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \right) \| z_k \|. \tag{17}$$

Supongamos que el máx $|z_i^{(k+1)}|$ se alcanza para algún $i = i_0$ así que $\| z_{k+1} \| = |z_{i_0}^{(k+1)}|$. De (17) para $i = i_0$ obtenemos

$$\left(1 - \sum_{j=1}^{i_0-1} \left| \frac{a_{i_0 j}}{a_{i_0 i_0}} \right| \right) \| z_{k+1} \| \leqslant q \left(1 - \sum_{j=1}^{i_0-1} \left| \frac{a_{i_0 j}}{a_{i_0 i_0}} \right| \right) \| z_k \|;$$

de aquí se deduce la estimación $\| z_{k+1} \| \leqslant q \| z_k \| \leqslant \ldots \leqslant q^{k+1} \| z_0 \|$. La afirmación está demostrada.

La condición (15) significa, que A es una matriz con predominancia diagonal. Para los ejemplos de aplicación del método de Seidel examinados en el punto 2 no se cumple la condición (15) ($q = 1$). En estos ejemplos el operador A es autoconjugado y definido positivo en H. Por eso en virtud del teorema 2 se puede solamente afirmar que el método converge en H_A. La estimación de la velocidad de convergencia en H_A será dado más abajo después del examen del esquema general de los métodos iterativos triangulares.

§ 2. Método de relajación superior

1. Esquema iterativo. Condiciones de convergencia suficientes. Para acelerar la convergencia del método de Seidel éste se modifica, introduciendo en el esquema iterativo un parámetro iterativo ω, de manera tal que

$$(\mathscr{D} + \omega L) \frac{y_{k+1} - y_k}{\omega} + A y_k = f, \quad k = 0, 1, \ldots, y_0 \in H, \tag{1}$$

donde, como antes, la matriz A está representada en forma de la suma

$$A = \mathscr{D} + L + U.$$

El método de Seidel corresponde al valor $\omega = 1$.

Comparando (1) con el tipo canónico de los esquemas iterativos de dos capas, encontramos que

$$B = \mathscr{D} + \omega L, \quad \tau_k \equiv \omega.$$

Como para el método de Seidel, para el método a examinar al operador B le corresponde una matriz triangular inferior, así que la introducción del parámetro ω no nos saca de la clase de los métodos iterativos triangulares. La pregunta sobre la elección del parámetro ω es nueva.

Si anotamos el esquema iterativo (1) según las componentes del vector y_{k+1} entonces obtendremos las fórmulas siguientes:

$$a_{ii}y_i^{(k+1)} = (1-\omega)\, a_{ii}y_i^{(k)} - \omega \sum_{j=1}^{i-1} a_{ij}y_j^{(k+1)} - \\ - \omega \sum_{j=i+1}^{M} a_{ij}y_j^{(k)} + \omega f_i \qquad (3)$$

para $i = 1, 2, \ldots, M$ (el $y_i^{(k+1)}$ hallado se sitúa en el lugar de $y_i^{(k)}$). La realización de un paso iterativo se lleva a efecto aproximadamente con los mismos gastos de operaciones aritméticas que en el método de Seidel.

El método iterativo (1) para $\omega > 1$ es llamado *método de relajación superior*, para $\omega = 1$ —*método de relajación completa* y para $\omega < 1$ — *método de relajación inferior*.

En el § 1 fue mostrado que el método de Seidel converge en H_A para el caso de un operador A autoconjugado y definido positivo. Para la convergencia del método de relajación, además de estas exigencias, se exige una condición complementaria sobre el parámetro iterativo ω. Formulemos las condiciones de convergencia suficientes del método de relajación.

TEOREMA 4. *Si el operador A es autoconjugado y definido positivo en H, y el parámetro ω satisface la condición $0 < \omega < 2$, entonces el método de relajación (1) converge en H_A.*

En efecto, del teorema 1 se desprende que es suficiente comprobar la desigualdad $\mathscr{D} + \omega L > 0{,}5\omega A$ para $\omega > 0$. Puesto que $A = A^* > 0$, entonces el operador $\mathscr{D}$ es autoconjugado y definido positivo en H y $U = L^*$. Por eso, utilizando la igualdad (14) del § 1, obtenemos

$$(\mathscr{D} + \omega L)\, x, x) = (1 - 0{,}5\omega)\, (\mathscr{D}x, x) + 0{,}5\omega\, ((\mathscr{D} + \\ + 2L)\, x, x) = (1 - 0{,}5\omega)\, (Dx, x) + 0{,}5\omega\, (Ax, x).$$

Para $\omega < 2$ de aquí se deduce la afirmación del teorema.

OBSERVACION. El teorema 4 es válido tanto para el método puntual de relajación, cuando en (3) los a_{ij} son los números, como para el método de relajación de bloques o vectorial, cuando en (3) los a_{ij} son las matrices de la dimensión correspondiente.

2. Planteamiento del problema sobre la elección del parámetro iterativo. El teorema 4 da condiciones de convergencia suficientes del método de relajación, dejando abierta la pregunta sobre la elección óptima del parámetro ω. La singularidad del proceso iterativo a examinar consiste en que el parámetro iterativo ω entra en el operador $B = \mathscr{D} + \omega L$, el cual es un operador autoconjugado en H. Con el caso no autoconjugado nosotros ya tuvimos que ver en el § 4 del cap. VI, donde fue examinado el método de iteración simple, para el cual el parámetro iterativo se eligió de diferentes condiciones, por ejemplo de la condición de mínimo de la norma del operador de transición de una iteración a otra. Aquí es necesario tener en cuenta la singularidad indicada más arriba del esquema iterativo. La elección del parámetro ω de la condición de mínimo de la norma en H_A del operador de transición de una iteración a otra será hecha en el § 3 de este capítulo, donde será examinado el esquema general de los métodos iterativos triangulares. En este punto el parámetro ω para el método de relajación se eligirá de la condición de mínimo del radio espectral del operador de transición de una iteración a otra.

Recordemos la definición del radio espectral de un operador

$$\rho(S) = \lim_{n\to\infty} \sqrt[n]{\| S^n \|} = \max_k |\lambda_k|, \tag{4}$$

donde λ_k son los valores propios del operador S. El radio espectral posee las siguientes propiedades:

$$\rho(S^n) = \rho^n(S), \quad \rho(S) \leqslant \| S \| \tag{5}$$

y $\rho(S) = \| S \|$, si S es un operador autoconjugado en H. De (5) para un operador arbitrario S obtenemos $\rho^n(S) = \rho(S^n) \leqslant \| S^n \|$. Por otra parte, de (4) para n suficientemente grande tendremos $\rho^n(S) \approx \| S^n \|$.

Pasemos ahora al planteamiento del problema sobre la elección óptima del parámetro ω para el esquema itera-

tivo (1). Obtendremos primeramente el problema para el error $z_k = y_k - u$. De (1) hallaremos

$$(\mathscr{D} + \omega L)\frac{z_{k+1} - z_k}{\omega} + Az_k = 0,\ k = 0,\ 1,\ \ldots,\ z_0 = y_0 - u$$

ó

$$z_{k+1} = Sz_k,\ k = 0,\ 1,\ \ldots,\ S = E - \omega(\mathscr{D} + \omega L)^{-1}A. \qquad (6)$$

Utilizando (6) expresaremos z_n mediante z_0:

$$z_n = S^n z_0, \quad \| z_n \| \leqslant \| S^n \| \| z_0 \|. \qquad (7)$$

El operador S es un operador no autoconjugado en H, dependiente del parámetro ω. Formulemos el problema sobre la elección óptima del parámetro ω de la forma siguiente: hallar ω de la condición de mínimo del radio espectral del operador S.

Se debe señalar que nosotros no minimizamos la norma del operador de resolución S^n, como se debía hacer en virtud de la estimación (7), sino minimizamos el radio espectral $\rho(S)$ del operador de transición S, para el cual tiene lugar la estimación $\rho^n(S) \leqslant \| S^n \|$. Sin embargo, en virtud de la igualdad aproximada $\rho^n(S) \approx \| S^n \|$ se puede esperar que para n suficientemente grande el procedimiento indicado de elección de ω resulte acertado.

La resolución del problema formulado más arriba es un problema complejo, sin embargo, bajo algunas suposiciones complementarias respecto al operador A este problema puede ser resuelto exitosamente.

SUPOSICION 1. El operador A es autoconjugado y definido positivo en H

$$H\ (U = L^*,\ \mathscr{D} = \mathscr{D}^* > 0).$$

SUPOSICIÓN 2. El operador A es tal, que para todo número complejo $z \neq 0$ los valores propios μ del problema generalizado de valores propios $\left(zL + \frac{1}{z}U\right)x - \mu\mathscr{D}x = 0$ no depende de z.

Utilizando estas suposiciones, demostremos la siguiente información, la cual no será útil en lo sucesivo.

LEMA 1. *Si el operador A satisface las suposiciones 1 y 2, entonces todos los valores propios del problema*

$$Ax - \lambda\mathscr{D}x = 0 \qquad (8)$$

son reales positivos y si λ es un valor propio, entonces $2 - \lambda$ también es un valor propio.

Realmente, la positividad y la realidad de los valores propios λ se deducen de la autoconjugación y definición positiva del operador A. A continuación, sea λ un valor propio del problema (8), es decir

$$Ax - \lambda \mathscr{D}x = (L + U)\,x - (\lambda - 1)\mathscr{D}x = 0, \quad x \neq 0.$$

En virtud de la suposición 2 tendrá lugar la igualdad

$$(-L - U)\,y - (\lambda - 1)\,\mathscr{D}y = 0 \text{ o } Ay - (2 - \lambda)\,\mathscr{D}y = 0.$$

De aquí se deduce la afirmación del lema.

Pasemos ahora a la resolución del problema sobre la elección óptima del parámetro ω. Para esto es necesario estimar el radio espectral del operador de transición $S = E - \omega(\mathscr{D} + \omega L)^{-1}A$, o sea estimar los valores propios μ del operador S:

$$Sx - \mu x = 0. \tag{9}$$

Consideraremos que son cumplidas las suposiciones 1 y 2. El siguiente lema establece la relación entre los valores propios μ del problema (9) y los valores propios λ del problema (8).

LEMA 2. Para $\omega \neq 1$ los valores propios de los problemas (8) y (9) están conectados por las relaciones

$$(\mu + \omega - 1)^2 = \omega^2 \mu\,(1 - \lambda)^2. \tag{10}$$

En efecto, sean μ y λ valores propios de los problemas (9) y (8). De la definición del operador S y del desarrollo $A = \mathscr{D} + L + U$ se deduce que (9) puede ser escrita en la forma

$$\frac{1-\mu-\omega}{\omega}\mathscr{D}x - (\mu L + U)\,x = 0, \quad x \neq 0. \tag{11}$$

Mostremos primeramente, que para $\omega \neq 1$ todos los μ son diferentes de cero. En efecto, supongamos, que $\mu = 0$. Entonces (11) toma la forma

$$\frac{1-\omega}{\omega}\mathscr{D}x - Ux = 0.$$

Ya que U es una matriz triangular superior, y $\mathscr{D}$ es una matriz diagonal (diagonal por bloques), la cual es definida positiva en virtud de la suposición 1, entonces la última igualdad puede tener lugar para $x \neq 0$ cuando y solamente cuando $w = 1$. Por consiguiente, hemos llegado a una contradicción, suponiendo que para $\omega \neq 1$ tenemos $\mu = 0$.

Dividiendo los miembros primero y segundo de (11) por $\sqrt{\mu}$, obtenemos

$$\frac{1-\mu-\omega}{\omega\sqrt{\mu}}\mathscr{D}x - \left(\sqrt{\mu}\,L + \frac{1}{\sqrt{\mu}}U\right)x = 0.$$

De aquí en virtud de la suposición 2 encontramos

$$\frac{1-\mu-\omega}{\omega\sqrt{\mu}}\mathcal{D}y-(L+U)\,y=0$$

o

$$Ay-\left(1+\frac{1-\mu-\omega}{\omega\sqrt{\mu}}\right)Dy=0.$$

Comparando esta igualdad con (8), obtenemos la relación

$$\frac{\mu+\omega-1}{\omega\sqrt{\mu}}=1-\lambda.$$

Con esto termina la demostración del lema 2.

OBSERVACIÓN. Durante la demostración del lema 2 no se utilizó la autoconjugación del operador A. La relación (10) tiene lugar también para el caso de todo operador A no autoconjugado bajo la suposición de no degeneración del operador $\mathcal{D}$.

Del lema 1 se deduce que los valores propios λ están distribuidos sobre el eje real simétricamente respecto al punto $\lambda=1$, y al mismo tiempo $\lambda\in[\lambda_{\text{mín}},\ 2-\lambda_{\text{mín}}]$. $\lambda_{\text{mín}}>0$. Por eso del lema 2 obtenemos que para $\omega\neq 1$ a cada $\lambda_i=1$ le corresponde $\mu_i=1-\omega$, a cada par λ_i y $2-\lambda_i$ le corresponde un par de μ_i no nulos, obtenidos por la solución de la ecuación (10) con $\lambda=\lambda_i$. Por lo tanto, todos los μ_i pueden ser hallados como las raíces de la ecuación cuadrática (10), en la cual en calidad de λ se toman todos los λ_i, distribuidos sobre el segmento $[\lambda_{\text{mín}},\ 1]$.

3. Estimación del radio espectral. Hallemos ahora el valor óptimo del parámetro ω y estimemos el radio espectral del operador S. Para esto investigaremos la ecuación (10):

$$\mu^2+[2(\omega-1)-\omega^2(1-\lambda)^2]\,\mu+(\omega-1)^2=0, \tag{12}$$

d n e $\lambda_{\text{mín}}\leqslant\lambda\leqslant 1$ y $0<\omega<2$.

Resolviendo la ecuación (12), hallaremos dos raíces

$$\mu_1(\lambda,\ \omega)=\left(\frac{\omega(1-\lambda)+\sqrt{\omega^2(1-\lambda)^2-4(\omega-1)}}{2}\right)^2,$$
$$\mu_2(\lambda,\ \omega)=\left(\frac{\omega(1-\lambda)-\sqrt{\omega^2(1-\lambda)^2-4(\omega-1)}}{2}\right)^2. \tag{13}$$

La investigación del discriminante de la ecuación (12) da, que para $\omega > \omega_0 > 1$, donde

$$\omega_0 = \frac{2}{1+\sqrt{\lambda_{\min}(2-\lambda_{\min})}} \in (1, 2), \tag{14}$$

las raíces μ_1 y μ_2 para cualquier $\lambda \in [\lambda_{\min}, 1]$ son complejas, y al mismo tiempo $|\mu_1| = |\mu_2| = \omega - 1$. Por eso el radio espectral del operador S para $\omega > \omega_0$ es igual a $\rho(S) = \omega - 1$ y crece según ω. Si $\omega = \omega_0$, entonces

$$\mu_1(\lambda_{\min}, \omega_0) = \mu_2(\lambda_{\min}, \omega_0) = \omega_0 - 1,$$

y para $\lambda_{\min} < \lambda \leqslant 1$ las raíces μ_1 y μ_2 de nuevo serán complejas y $|\mu_1| = |\mu_2| = \omega_0 - 1$. Por consiguiente, en la región $\omega \geqslant \omega_0$ el valor $\omega = \omega_0$ es óptimo, al cual le corresponde $\rho(S) = \omega_0 - 1$.

Sea ahora $1 < \omega < \omega_0$. Investiguemos el comportamiento de las raíces μ_1 y μ_2, definidas por la fórmula (13), como funciones de la variable λ para ω fijo.

Si λ pertenece al segmento $[\lambda_{\min}, \lambda_0]$,

$$\lambda_{\min} \leqslant \lambda \leqslant \lambda_0 = 1 - 2\frac{\sqrt{\omega-1}}{\omega} < 1,$$

entonces el discriminante $\omega^2(1-\lambda)^2 - 4(\omega-1)$ no es negativo y, por consiguiente, las raíces μ_1 y μ_2 son reales y al mismo tiempo la raíz μ_1 es máxima.

Mostremos, que $\mu_1(\lambda, \omega)$ es una función decreciente de λ en el segmento $[\lambda_{\min}, \lambda_0]$. En efecto, diferenciando (12) según λ y teniendo en cuenta (13), obtenemos

$$\frac{\partial \mu_1}{\partial \lambda} = -\frac{2\omega\mu_1}{\sqrt{\omega^2(1-\lambda)^2 - 4(\omega-1)}} > 0.$$

Por lo tanto, la raíz $\mu_1(\lambda, \omega)$ para $1 < \omega < \omega_0$ decrece al variar λ de $\lambda_{\min}$ a λ_0, tomando los siguientes valores:

$$\mu_1(\lambda_{\min}, \omega) = \left(\frac{\omega(1-\lambda_{\min}) + \sqrt{\omega^2(1-\lambda_{\min})^2 - 4(\omega-1)}}{2}\right)^2,$$

$$\mu_1(\lambda_0, \omega) = \omega - 1.$$

Así sucesivamente, si λ cambia de λ_0 hasta 1, entonces las raíces μ_1 y μ_2 son complejas e iguales en módulo: $|\mu_1| = |\mu_2| = \omega - 1$. Por consiguiente, si $1 < \omega < \omega_0$, entonces

$$\rho(S) = \mu_1(\lambda_{\min}, \omega) = \left(\frac{\omega(1-\lambda_{\min} + \sqrt{\omega^2(1-\lambda_{\min})^2 - 4(\omega-1)}}{2}\right)^2. \tag{15}$$

Si $\omega < 1$, entonces todas las raíces de la ecuación (12) son reales y la raíz máxima es μ_1, cuyos valores decrecen al variar λ de $\lambda_{\min}$ hasta 1. Por lo tanto, para $\omega < 1$ el radio espectral del operador S se determina por la fórmula (15). Puesto que para $\omega = 1$ los μ_k no

nulos satisfacen la ecuación (12), entonces (15) tiene lugar tambien para $\omega = 1$.

De esta forma, si $0 < \omega < \omega_0$, entonces el radio espectral del operador S se determina por la fórmula (15). Mostremos que $\mu_1(\lambda_{min}, \omega)$ decrece según ω en el intervalo $0 < \omega < \omega_0$.

En efecto, como para $\omega < \omega_0$, la raíz μ_1 decrece según λ para $\lambda \leqslant \lambda_0$, y $\mu_1(0, \omega) = 1$, entonces $\mu_1(\lambda_{min}, \omega) < 1$.

A continuación, de (15) obtenemos

$$\frac{\partial \mu_1(\lambda_{min}, \omega)}{\partial \omega} = \sqrt{\mu_1}\left(1-\lambda_{min}+\frac{\omega(1-\lambda_{min})^2-2}{\sqrt{\omega^2(1-\lambda_{min})^2-4(\omega-1)}}\right)=$$

$$=\frac{\sqrt{\mu_1}}{\omega}\,\frac{[\omega^2(1-\lambda_{min})^2-2(\omega-1)+(1-\lambda_{min})\omega\sqrt{\omega^2(1-\lambda_{min})^2-4(\omega-1)}-2]}{\sqrt{\omega^2(1-\lambda_{min})^2-4(\omega-1)}}.$$

Sustituyendo aquí (13), definitivamente hallaremos

$$\frac{\partial \mu_1}{\partial \omega} = \frac{2\sqrt{\mu_1}(\mu_1-1)}{\omega\sqrt{\omega^2(1-\lambda_{min})^2-4(\omega-1)}} < 0.$$

La afirmación está demostrada. Por lo tanto, en la región $\omega \leqslant \omega_0$ el valor $\omega = \omega_0$ es óptimo y a él le corresponde

$$\rho(S) = \omega_0 - 1 = \frac{1-\sqrt{\lambda_{min}(2-\lambda_{min})}}{1+\sqrt{\lambda_{min}(2-\lambda_{min})}} = \left(\frac{1-\sqrt{\eta}}{1+\sqrt{\eta}}\right)^2, \quad \eta = \frac{\lambda_{min}}{2-\lambda_{min}}.$$

Notemos, que de las investigaciones realizadas más arriba se desprende que si δ es la estimación para λ_{min} por debajo, es decir, $\delta \leqslant \lambda_{min}$, y ω es elegido según la fórmula (14) cambiando λ_{min} por δ, entonces $\omega_0 \leqslant \omega$,

$$\rho(S) \leqslant \left(\frac{1-\sqrt{\eta}}{1+\sqrt{\eta}}\right)^2, \quad \eta = \frac{\delta}{2-\delta}.$$

Así, hemos demostrado el siguiente

TEOREMA 5. *Supongamos que estén cumplidas las suposiciones 1 y 2 y δ es la constante de la desigualdad*

$$\delta_{\mathcal{D}} \leqslant A, \quad \delta > 0. \tag{16}$$

Entonces para el radio espectral del operador de transición S del esquema iterativo (1) *siendo óptimo el valor del parámetro ω,*

$$\omega = \omega_0 = \frac{2}{1+\sqrt{\delta(2-\delta)}}; \tag{17}$$

es válida la estimación

$$\rho(S) \leqslant \left(\frac{1-\sqrt{\eta}}{1+\sqrt{\eta}}\right)^2, \quad \eta = \frac{\delta}{2-\delta}, \tag{18}$$

y al mismo tiempo, si en (16) *se alcanza la igualdad, entonces ella tiene lugar también en la fórmula* (18).

El método iterativo (1), (17) es el *método de relajación superior*, ya que $\omega_0 > 1$.

4. **Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo.** Examinemos la aplicación del método de relajación superior para encontrar una solución aproximada del problema de Dirichlet de diferencias para la ecuación de Poisson, prefijada sobre la red rectangular $\overline{\omega} = \{x_{ij} = (ih_1,\ jh_2) \in \overline{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1, 2\}$ en el rectángulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$:

$$\Lambda y = \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha x_\alpha} = -\varphi(x),\ x \in \omega,\ y(x) = g(x),\ x \in \gamma. \tag{19}$$

El operador A en el espacio H de las funciones reticulares, definidas sobre ω con el producto escalar

$$(u,\ v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2$$

se define por el procedimiento usual: $Ay = -\Lambda \mathring{y}$, $y \in H$, $\mathring{y} \in \mathring{H}$. Como nosotros ya sabemos, el operador A, correspondiente al problema (19), es autoconjugado y definido positivo en H. Por consiguiente, se cumple la suposición 1.

Examinemos primeramente el *método puntual de relajación superior*. Si las incógnitas son ordenadas según las filas de la red ω, entonces el esquema de diferencias (19) puede ser escrito en la forma del siguiente sistema de ecuaciones algebraicas:

$$-\frac{1}{h_1^2} y(i-1,\ j) - \frac{1}{h_2^2} y(i,\ j-1) + \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y(i,\ j) - \\ - \frac{1}{h_1^2} y(i+1,\ j) - \frac{1}{h_2^2} y(i,\ j+1) = \varphi(i,\ j)$$

para $i = 1, 2, \ldots, N_1 - 1$, $j = 1, 2, \ldots, N_2 - 1$ e $y(x) = g(x)$, $x \in \gamma$.

A esta escritura del operador A le corresponde la representación de A en forma de la suma $A = \mathscr{D} + L + U$, donde

$$\mathscr{D}y = \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y,$$

$$Ly(i, j) = -\frac{1}{h_1^2}\overset{\circ}{y}(i-1, j) - \frac{1}{h_2^2}\overset{\circ}{y}(i, j-1),$$

$$Uy(i, j) = -\frac{1}{h_1^2}\overset{\circ}{y}(i+1, j) - \frac{1}{h_2^2}\overset{\circ}{y}(i, j+1).$$

Para el sistema a examinar el método puntual de relajación superior, en correspondencia con la fórmula (3), tendrá la forma siguiente:

$$\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y_{k+1}(i, j) = (1-\omega)\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y_k(i, j) +$$
$$+ \omega\left[\frac{1}{h_1^2} y_{k+1}(i-1, j) + \frac{1}{h_2^2} y_{k+1}(i, j-1) +\right.$$
$$\left. + \frac{1}{h_1^2} y_k(i+1, j) + \frac{1}{h_2^2} y_k(i, j+1) + \varphi(i, j)\right]$$

para $i = 1, 2, \ldots, N_1 - 1$, $j = 1, 2, \ldots, N_2 - 1$, y al mismo tiempo $y_k(x) = g(x)$ siendo $x \in \gamma$ para cualquier $k \geqslant 0$.

Los cálculos, al igual que en el método de Seidel, se comienzan por el punto $i = 1$, $j = 1$ y se continúan o bien por las filas o por las columnas de la red ω. El $y_{k+1}(i, j)$ hallado se sitúa en el lugar de $y_k(i, j)$.

Demostremos ahora que para el ejemplo a examinar se cumple la suposición 2. Para esto es necesario mostrar que para cualquier número complejo $z \neq 0$ los valores propios μ del problema

$$z\left(\frac{1}{h_1^2} y(i-1, j) + \frac{1}{h^2} y(i, j-1)\right) + \frac{1}{z}\left(\frac{1}{h_1^2} y(i+1, j) +\right.$$
$$\left. + \frac{1}{h_2^2} y(i, j+1)\right) + \mu\left[\frac{2}{h_1^2} + \frac{2}{h_2^2}\right] y(i, j) = 0,$$
$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1,$$
$$y(x) = 0, \quad x \in \gamma$$

no dependen de z.

En efecto, poniendo aquí

$$y(i, j) = z^{i+j} v(i, j), \quad 0 \leqslant i \leqslant N_1, \quad 0 \leqslant j \leqslant N_2,$$

obtendremos

$$\frac{1}{h_1^2} v(i-1,\ j) + \frac{1}{h^2} v(i,\ j-1) + \frac{1}{h_1^2} v(i+1,\ j) + \\ + \frac{1}{h_2^2} v(i,\ j+1) + \mu \left[\frac{2}{h_1^2} + \frac{2}{h_2^2}\right] v(i,\ j) = 0,$$

$$1 \leqslant i \leqslant N_1 - 1,\ 1 \leqslant j \leqslant N_2 - 1,\ v(x) = 0,\ x \in \gamma.$$

Por consiguiente, μ no depende de z.

Queda por hallar el valor óptimo del parámetro ω. Para esto es necesario hallar o estimar por debajo el valor propio mínimo del problema (8), el cual para este caso se escribe en la forma

$$y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + \lambda \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y = 0,\quad x \in \omega,\quad y(x) = 0,\quad x \in \gamma.$$

Puesto que los valores propios del operador de Laplace de diferencias $\overset{\circ}{\Lambda} y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2}$ son conocidos

$$\overset{\circ}{\lambda}_k = \frac{4}{h_1^2} \operatorname{sen}^2 \frac{k_1 \pi h_1}{2l_1} + \frac{4}{h_2^2} \operatorname{sen}^2 \frac{k_2 \pi h_2}{2l_2},$$

$$k_\alpha = 1,\ 2,\ \ldots,\ N_\alpha - 1,$$

entonces

$$\lambda_k = \overset{\circ}{\lambda}_k \Big/ \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) = \frac{2h_2^2}{h_1^2 + h_2^2} \operatorname{sen}^2 \frac{k_1 \pi h_1}{2l_1} + \\ + \frac{2h_1^2}{h_1^2 + h_2^2} \operatorname{sen}^2 \frac{k_2 \pi h_2}{2l_2}.$$

Por lo tanto,

$$\lambda_{\min} = \frac{2h_2^2}{h_1^2 + h_2^2} \operatorname{sen}^2 \frac{\pi h_1}{2l_1} + \frac{2h_1^2}{h_1^2 + h_2^2} \operatorname{sen}^2 \frac{\pi h_2}{2l_2},$$

y el parámetro ω_0 se encuentra por la fórmula (14). En el caso particular, cuando G es un cuadrado con lado l ($l_1 = l_2 = l$) y la red es cuadrada ($N_1 = N_2 = N$), tenemos

$$\lambda_{\min} = 2 \operatorname{sen}^2 \frac{\pi}{2N},\quad \omega_0 = \frac{2}{1 + \operatorname{sen} \frac{\pi}{N}},\quad \eta = \operatorname{tg}^2 \frac{\pi}{2N},$$

$$\rho(S) = \frac{1 - \operatorname{sen} \frac{\pi}{N}}{1 + \operatorname{sen} \frac{\pi}{N}} \approx 1 - \frac{2\pi}{N}.$$

Notemos, que el radio espectral del operador de transición correspondiente al método de Seidel puntual se estima por la fórmula (15), en la cual se debe poner $\omega = 1$. Esto da $\rho(S) = (1 - \lambda_{\min})^2 = \cos^2 \frac{\pi}{N}$, lo cual es considerablemente peor, que para el método de relajación superior.

Examinemos ahora el método de bloques de relajación superior. Si a un bloque se le unen las incógnitas $y(i, j)$ sobre la j-ésima fila de la red, entonces a la escritura por bloques del operador A le corresponde la representación siguiente $A = \mathscr{D} + L + U$, donde[1]

$$\mathscr{D}y = -\frac{1}{h_1^2}\mathring{y}(i-1, j) + \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right)\mathring{y}(i, j) - \frac{1}{h_1^2}\mathring{y}(i+1, j),$$

$$Ly(i, j) = -\frac{1}{h_2^2}\mathring{y}(i, j-1), \quad Uy(i, j) = -\frac{1}{h_2^2}\mathring{y}(i, j+1).$$

Las fórmulas de cálculo para el método de bloques de relajación superior tienen la forma

$$-\frac{1}{h_1^2}y_{k+1}(i-1, j) + \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right)y_{k+1}(i, j) -$$

$$-\frac{1}{h_2^2}y_{k+1}(i+1, j) = (1-\omega)\left(-\frac{1}{h_1^2}y_k(i-1, j) + \right.$$

$$+\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right)y_k(i, j) - \frac{1}{h_2^2}y_k(i+1, j)\Big) +$$

$$+\omega\left(\frac{1}{h_2^2}y_{k+1}(i, j-1) + \frac{1}{h_2^2}y_k(i, j+1) + \varphi(i, j)\right),$$

$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1,$$

y al mismo tiempo $y_k(x) = g(x)$, $x \in \gamma$ para todos los $k \geqslant 0$. Para encontrar y_{k+1} sobre la j-ésima fila es necesario resolver, por ejemplo, por el método de factorización un problema de contorno tripuntual.

Mostremos que para el ejemplo examinado se cumple la suposición 2, es decir, los valores propios μ del problema

$$z\frac{1}{h_2^2}y(i, j-1) + \frac{1}{z}\frac{1}{h_2^2}y(i, j+1) + \mu\left(-\frac{1}{h_1^2}y(i-1, j) + \right.$$

$$+\left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right)y(i, j) - \frac{1}{h_2^2}y(i+1, j)\Big) = 0,$$

$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1, \quad y(x) = 0, \quad x \in \gamma$$

no dependen de z. Esto se establece fácilmente con ayuda del cambio $y(i, j) = z^j v(i, j)$, $0 \leqslant i \leqslant N_1$, $0 \leqslant j \leqslant N_2$.

Hallaremos ahora el valor óptimo del parámetro ω. El problema correspondiente (8) tiene la forma

$$\begin{aligned} &y_{\bar{x}_1x_1} + y_{\bar{x}_2x_2} + \lambda\left(\frac{2}{h_2^2}\, y - y_{\bar{x}_1x_1}\right) = 0, \quad x \in \omega, \\ &y(x) = 0, \quad x \in \gamma. \end{aligned} \tag{20}$$

No es difícil comprobar, que las funciones propias del problema (20) son

$$y_k(x) = \operatorname{sen} \frac{k_1 \pi x_1}{l_1} \operatorname{sen} \frac{k_2 \pi x_2}{l_2}. \tag{21}$$

Sustituyendo (21) en (20), hallaremos

$$\lambda_k = \frac{\overset{\circ}{\lambda}_{k_1} + \overset{\circ}{\lambda}_{k_2}}{\frac{2}{h_2^2} + \overset{\circ}{\lambda}_{k_1}}, \quad k_\alpha = 1, 2, \ldots, N_\alpha - 1, \quad k = (k_1, k_2),$$

donde

$$\overset{\circ}{\lambda}_{k_\alpha} = \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{k_\alpha \pi h_\alpha}{2l_\alpha}, \quad k_\alpha = 1, 2, \ldots, N_\alpha - 1, \alpha = 1, 2.$$

De aquí obtendremos

$$\lambda_{\min} = \frac{2h_2^2 \operatorname{sen}^2 \frac{\pi h_1}{2l_1} + 2h_1^2 \operatorname{sen}^2 \frac{\pi h_2}{2l_2}}{2h_2^2 \operatorname{sen}^2 \frac{\pi h_1}{2l_1} + h_1^2}.$$

Para el caso particular examinado más arriba tendremos

$$\lambda_{\min} = \frac{4 \operatorname{sen}^2 \frac{\pi}{2N}}{1 + 2 \operatorname{sen}^2 \frac{\pi}{2N}}, \quad \omega_0 = \frac{2 + 4 \operatorname{sen}^2 \frac{\pi}{2N}}{\left(1 + \sqrt{2} \operatorname{sen} \frac{\pi}{2N}\right)^2},$$

$$\eta = 2 \operatorname{sen}^2 \frac{\pi}{2N}, \quad \rho(S) = \left(\frac{1 - \sqrt{2} \operatorname{sen} \frac{\pi}{2N}}{1 + \sqrt{2} \operatorname{sen} \frac{\pi}{2N}}\right)^2 \approx 1 - 2\sqrt{2}\frac{\pi}{N}.$$

Comparando las estimaciones del radio espectral de los métodos de relajación superior de bloques y puntual, encontramos, que el método de bloques convergerá en $\sqrt{2}$ veces más rápido que el método puntual. Por otra parte, el método de bloques exige un mayor número de operaciones aritméticas gastadas en la realización de un paso iterativo que el método puntual.

Como conclusión exponemos el número de iteraciones para el método puntual de relajación superior en dependencia del número de nodos N por una dirección para $\varepsilon = 10^{-4}$. En calidad de problema modelo tomemos el esquema de diferencias (19) sobre la red cuadrada con $N_1 = N_2 = N$ y $\varphi(x) \equiv 0$, $g(x) \equiv 0$. Eligiremos la aproximación inicial $y_0(x)$ de la siguiente forma: $y_0(x) = 1$, $x \in \omega$, $y_0(x) = 0$, $x \in \gamma$.

Terminaremos el proceso de iteraciones, si se cumple la condición

$$\| z_n \|_A \leqslant \varepsilon \| z_0 \|_A. \tag{22}$$

De la teoría del método se deduce, que para el error z_n tiene lugar la estimación $\| z_n \|_A \leqslant \| S^n \|_A \| z_0 \|_A$, puesto que el radio espectral del operador es menor o igual que cualquier norma del operador, entonces $\rho^n(S) \leqslant \| S^n \|_A$. Por lo tanto, la condición $\rho^n(S) \leqslant \varepsilon$ no se puede emplear para la estimación del número exigido de iteraciones.

Citemos el número de iteraciones n, determinado de la condición (22), y para comparar hallaremos el número de iteraciones n^*, el cual se deduce de la desigualdad $\rho^n(S) \leqslant \varepsilon$:

$N = 32$	$n = 65$	$n^* = 47$
$N = 64$	$n = 128$	$n^* = 94$
$N = 128$	$n = 257$	$n^* = 187$

La comparación del número de iteraciones para el método de relajación superior y el método explícito de Chebishev, examinado para el problema (19) en el punto 1 del § 5 del cap. VI, muestra que el método de relajación superior exige aproximadamente en 1,6 veces menos iteraciones que el método explícito de Chebishev. El número de operaciones aritméticas gastadas en una iteración es prácticamente igual en estos métodos.

5. Problema de Dirichlet de diferencias para una ecuación elíptica con coeficientes variables. Examinemos ahora la aplicación del método de relajación superior para encontrar una solución aproximada del problema de Dirichlet de diferencias para la ecuación con coeficientes variables en un rectángulo

$$\begin{aligned} &\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha(x)\, y_{\bar{x}_\alpha})_{x_\alpha} - d(x)\, y = -\varphi(x), \quad x \in \omega, \\ &y(x) = g(x), \quad x \in \gamma, \end{aligned} \tag{23}$$

considerando, que se cumplen las condiciones siguientes:

$$0 < c_1 \leqslant a_\alpha(x) \leqslant c_2, \quad x \in \bar{\omega}, \ \alpha = 1, 2,$$
$$0 \leqslant d_1 \leqslant d(x) \leqslant d_2, \quad x \in \omega. \tag{24}$$

Para el problema (23) el método puntual de relajación superior para ordenar las incógnitas según las filas de la red ω se describe por la fórmula

$$b(i, j) y_{k+1}(i, j) = (1-\omega) b(i, j) y_k(i, j) +$$
$$+ \omega \Big[\frac{a_1(i, j)}{h_1^2} y_{k+1}(i-1, j) + \frac{a_2(i, j)}{h_2^2} y_{k+1}(i, j-1) +$$
$$+ \frac{a_1(i+1, j)}{h_1^2} y_k(i+1, j) + \frac{a_2(i, j+1)}{h_2^2} y_k(i, j+1) + \varphi(i, j) \Big],$$
$$1 \leqslant i \leqslant N_1 - 1, \ 1 \leqslant j \leqslant N_2 - 1, \tag{25}$$

donde

$$b(i, j) = \frac{a_1(i, j) + a_1(i+1, j)}{h_1^2} + \frac{a_2(i, j) + a_2(i, j+1)}{h_2^2} + d(i, j)$$

e $y_k(x) = g(x)$, $x \in \gamma$ para cualquier $k \geqslant 0$.

Para el ejemplo a examinar los operadores $\mathscr{D}$, L y U se definen de la siguiente forma:
$\mathscr{D}y = by$,

$$Ly(i, j) = -\frac{a_1(i, j)}{h_1^2} \overset{\circ}{y}(i-1, j) - \frac{a_2(i, j)}{h_2^2} \overset{\circ}{y}(i, j-1),$$

$$Uy(i, j) = -\frac{a_1(i+1, j)}{h_1^2} \overset{\circ}{y}(i+1, j) - \frac{a_2(i, j+1)}{h_2^2} \overset{\circ}{y}(i, j+1).$$

Las suposiciones 1 y 2 se cumplen, lo cual se demuestra igual que para el ejemplo del punto 4.

Para hallar el parámetro ω, es necesario estimar la constante δ en la desigualdad $A \geqslant \delta \mathscr{D}$. Este problema fue resuelto anteriormente en el punto 3 del § 5 del cap. VI, donde se examinó el método de Chebishev implícito más simple para el problema de diferencias (23). Exponemos la estimación para δ:

$$\delta = \min_{0 < x_2 < l_2} \frac{1}{\varkappa_1(x_2)} + \min_{0 < x_1 < l_1} \frac{1}{\varkappa_2(x_1)},$$

donde $\varkappa_\alpha(x_\beta) = \max\limits_{0<x_\alpha<l_\alpha} v^\alpha(x)$, $\beta = 3-\alpha$, $\alpha = 1, 2$ y $v^\alpha(x)$ es la solución del siguiente problema de contorno tripuntual:

$$(a_\alpha v^\alpha_{\bar{x}_\alpha})_{x_\alpha} - \frac{1}{2}\, dv^\alpha = -b(x),\ h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,$$

$$v_\alpha(x) = 0, \quad x_\alpha = 0,\ l_\alpha,$$

$$h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta, \quad \beta = 3-\alpha, \quad \alpha = 1, 2.$$

El parámetro iterativo ω se encuentra mediante la fórmula (17):

$$\omega = \omega_0 = \frac{2}{1+\sqrt{\delta(2-\delta)}}.$$

Para comparar el método descrito de relajación superior con el método de Chebishev explícito más simple, examinado en el punto 3 del § 5 del cap. VI, citemos el número de iteraciones del método de relajación superior para el siguiente ejemplo modelo. Sea el esquema de diferencias (23) dado sobre una red cuadrada con $N_1 = N_2 = N$ y $\varphi(x) = 0$, $g(x) = 0$. Eligiremos los coeficientes $a_1(x)$, $a_2(x)$ y $d(x)$ de la siguiente forma:

$$a_1(x) = 1 + c\,[(x_1 - 0{,}5)^2 + (x_2 - 0{,}5)^2],$$

$$a_2(x) = 1 + c\,[0{,}5 - (x_1 - 0{,}5)^2 - (x_2 - 0{,}5)^2],$$

$$d(x) \equiv 0, \quad c > 0.$$

Con esto en las desigualdades (24) $c_1 = 1$, $c_2 = 1 + 0{,}5c$, $d_1 = d_2 = 0$. La aproximación inicial para el método iterativo de relajación superior (25) se elige de la siguiente forma: $y_0(x) = 1$, $x \in \omega$, $y_0(x) = 0$, $x \in \gamma$, y el proceso de las iteraciones lo terminaremos al cumplirse la condición (22).

Tabla 9

c_1/c_2	2	8	32	128	512
$N=32$	65	81	95	96	98
$N=64$	129	164	192	193	195

En la tabla 9 se dan los números de iteraciones para el método de relajación en función de la relación c_2/c_1 y del número de nodos N por una dirección para $\varepsilon = 10^{-4}$. Para

el caso, cuando $a_\alpha(x) \equiv 1$ y $d(x) \equiv 0$, el número de iteraciones del método de relajación superior se muestra en el punto 4 del presente párrafo.

De la tabla 9 se deduce, que el número de iteraciones del método de relajación superior para el ejemplo modelo es aproximadamente en dos veces menor que el número de iteraciones del método de Chebishev implícito más simple. Ya que el número de operaciones aritméticas, gastadas en la realización de un paso iterativo para los métodos indicados es igual, entonces el método de relajación superior es aproximadamente en dos veces más efectivo que el método de Chebishev implícito más simple.

§ 3. Métodos triangulares

1. **Esquema iterativo.** En los §§ 1, 2 fueron estudiados dos métodos: el método de Seidel y el método de relajación. Estos métodos pertenecen a la clase de los métodos implícitos de dos capas, en los cuales al operador B le corresponde una matriz triangular o triangular por bloques. En la forma canónica el esquema iterativo de los métodos tiene la forma siguiente:

$$(\mathscr{D} + \omega L)\frac{y_{k+1} - y_k}{\omega} + Ay_k = f,\ k = 0,\ 1,\ \ldots,\ y_0 \in H, \qquad (1)$$

donde $\mathscr{D}$ y L son los operadores del desarrollo de A en suma de una matriz diagonal, una triangular superior y otra triangular inferior

$$A = \mathscr{D} + L + U.$$

Al método de Seidel le corresponde el valor del parámetro $\omega = 1$.

Para el caso de un operador A autoconjugado y definido positivo en H la condición suficiente de convergencia del método iterativo (1) en H_A tiene la forma

$$0 < \omega < 2. \qquad (3)$$

En el § 2 hemos examinado el problema de la elección óptima del parámetro iterativo ω. Considerando que están cumplidas las suposiciones 1 y 2 y la información a priori está dada en forma de la constante δ de la desigualdad

$$\delta\mathscr{D} \leqslant A, \quad \delta > 0, \qquad (4)$$

hemos demostrado que el valor óptimo de ω, para el cual se minimiza el radio espectral del operador de transición S

del esquema (1), se determina por la fórmula

$$\omega = \omega_0 = \frac{2}{1+\sqrt{\delta(2-\delta)}}. \qquad (5)$$

En los puntos 4 y 5 del § 2 han sido examinados ejemplos de problemas para los cuales están cumplidas las suposiciones 1 y 2. Estas suposiciones son cumplidas también para problemas más complejos, por ejemplo para el esquema de diferencias pentapuntual, el cual aproxima el problema de Dirichlet para una ecuación elíptica con coeficientes variables sobre una red no uniforme en una región arbitraria.

Existen, sin embargo, ejemplos de problemas, para los cuales no se cumple la suposición 2. A ellos pertenecen el problema de Dirichlet de diferencias para una ecuación elíptica con derivadas mixtas, el problema de Dirichlet de diferencias de orden elevado de exactitud y otros.

La no universalidad del procedimiento de elección del parámetro iterativo ω y la ausencia de estimaciones de la velocidad de convergencia del método en alguna norma son las insuficiencias fundamentales de la teoría desarrollada en el § 2.

En el presente párrafo será examinado el esquema general de los métodos iterativos triangulares, para los cuales el parámetro iterativo ω se elige de la condición de minimizar la norma del operador de transición en H_A. Aquí será hallada la estimación de la velocidad de convergencia del método en H_A bajo la suposición de autoconjugación y definición positiva del operador A.

Comenzaremos el examen de los métodos triangulares por la transformación del esquema iterativo (1). Introduzcamos los operadores R_1 y R_2 de la siguiente forma:

$$R_1 = \frac{1}{2}\mathscr{D} + L, \quad R_2 = \frac{1}{2}\mathscr{D} + U.$$

Entonces el desarrollo (2) tendrá la forma

$$A = R_1 + R_2, \qquad (6)$$

y si A es un operador autoconjugado en H, entonces los operadores R_1 y R_2 son conjugados uno a otro

$$R_1 = R_2^*. \qquad (7)$$

Sustituyendo $L = R_1 - \frac{1}{2}\mathscr{D}$ en (1) y designando

$$\tau = 2\omega/(2-\omega), \qquad (8)$$

escribiremos el esquema iterativo (1) en la forma equivalente

$$(\mathscr{D} + \tau R_1)\frac{y_{k+1} - y_k}{\tau} + Ay_k = f,\ k = 0,\ 1,\ \ldots,\ y_0 \in H, \quad (9)$$

al mismo tiempo en virtud de (3) y (8), $\tau > 0$.

El esquema (9) se puede examinar independientemente del esquema (1). Precisamente, sea representado por la fórmula (6) en forma de la suma de operadores R_1 y R_2 conjugados uno a otro el operador A autoconjugado en H, y $\mathscr{D}$, un operador arbitrario autoconjugado y definido positivo en H. Al esquema iterativo (9) lo llamaremos *forma canónica de los métodos iterativos triangulares*. Nosotros conservamos el nombre de métodos triangulares aún en el caso, cuando las matrices, correspondientes a los operadores R_1 y R_2, no son triangulares, y la matriz, correspondiente al operador $\mathscr{D}$, no es una matriz diagonal.

Del teorema 1 se sigue que para un operador definido positivo A el método iterativo (9) con $\tau > 0$ converge en H_A. En efecto, para esto es suficiente establecer la validez de la desigualdad $\mathscr{D} + \tau R_1 > 0{,}5\tau A$. De (7) obtenemos

$$(Ax,\ x) = (R_1 x,\ x) + (R_2 x,\ x) = 2\,(R_1 x,\ x) = 2\,(R_2 x,\ x) \quad (10)$$

y, por consiguiente,

$$((\mathscr{D} + \tau R_1)\,x,\ x) = (\mathscr{D}x,\ x) + 0{,}5\tau\,(Ax,\ x) > 0{,}5\tau\,(Ax,\ x),$$

lo que se exigía demostrar.

Como conclusión señalemos, que al método de Seidel en el esquema (9) le corresponde el valor $\tau = 2$, y al método de relajación superior, $\tau = 2/\sqrt{\delta\,(2 - \delta)}$.

2. **Estimación de la velocidad de convergencia.** Estimemos ahora la velocidad de convergencia del esquema iterativo (9) en H_A, suponiendo que A es un operador autoconjugado y definido positivo en H.

El paso en (9) al error $z_k = y_k - u$ da un esquema homogéneo para z_k

$$B\frac{z_{k+1} - z_k}{\tau} + Az_k = 0,\ k = 0,\ 1.\ \ldots,\ z_0 = y_0 - u,\ B = \mathscr{D} + \tau R_1$$

de donde obtenemos

$$z_{k+1} = Sz_k, \quad k = 0, 1, \ldots, S = E - \tau B^{-1}A, \quad (11)$$
$$\| z_{k+1} \|_A \leqslant \| S \|_A \| z_k \|_A.$$

Estimemos la norma del operador de transición S en H_A. De la definición de la norma del operador obtendremos

$$\| S \|_A^2 = \sup_{x \neq 0} \frac{(ASx, Sx)}{(Ax, x)} = \sup_{x \neq 0} \left[1 - 2\tau \frac{(B^{-1}Ax, Ax)}{(Ax, x)} + \right.$$
$$\left. + \tau^2 \frac{(AB^{-1}Ax, B^{-1}Ax)}{(Ax, x)} \right]. \quad (12)$$

Transformemos la expresión que se encuentra entre corchetes. Utilizando (10) y la definición del operador B, obtendremos

$$(By, y) = (\mathscr{D}y, y) + \tau (R_1 y, y) = (\mathscr{D}y, y) +$$
$$+ 0{,}5\tau (Ay, y).$$

De aquí hallaremos $\tau^2 (Ay, y) = 2\tau (By, y) - 2\tau (\mathscr{D}y, y)$ o después de la sustitución $y = B^{-1}Ax$

$$\tau^2 (AB^{-1}Ax, B^{-1}Ax) = 2\tau (B^{-1}Ax, Ax) -$$
$$- 2\tau (\mathscr{D}B^{-1}Ax, B^{-1}Ax).$$

Sustituyendo esta expresión en (12), tendremos

$$\| S \|_A^2 = \sup_{x \neq 0} \left[1 - 2\tau \frac{(\mathscr{D}B^{-1}Ax, B^{-1}Ax)}{(Ax, x)} \right].$$

Producimos las siguientes transformaciones. Poniendo $x = (B^*)^{-1}\mathscr{D}^{1/2}y$, obtendremos

$$\frac{(\mathscr{D}B^{-1}Ax, B^{-1}Ax)}{(Ax, x)} = \frac{(Cy, Cy)}{(Cy, y)}, \quad C = \mathscr{D}^{1/2}B^{-1}A(B^*)^{-1}\mathscr{D}^{1/2}.$$

Ya que el operador C es autoconjugado y definido positivo en H, entonces, poniendo $y = C^{-1/2}\mathscr{D}^{-1/2}B^*v$, hallaremos

$$\frac{(\mathscr{D}B^{-1}Ax, B^{-1}Ax)}{(Ax, x)} = \frac{(Av, v)}{(B\mathscr{D}^{-1}B^*v, v)}.$$

Así, definitivamente tendremos

$$\| S \|_A^2 = \sup_{v \neq 0} \left[1 - 2\tau \frac{(Av, v)}{(B\mathscr{D}^{-1}B^*v, v)} \right].$$

De aquí obtendremos, que si γ_1 es la magntud de la desigualdad

$$\gamma_1 B\mathscr{D}^{-1}B^* \leqslant A, \tag{13}$$

entonces

$$\| S \|_A \leqslant (1 - 2\tau\gamma_1)^{1/2}. \tag{14}$$

Como γ_1 depende del parámetro τ, entonces el valor óptimo para τ se podrá hallar, obteniendo una expresión para γ_1 bajo ciertas suposiciones complementarias respecto a los operadores $\mathscr{D}$, R_1 y R_2.

3. **Elección del parámetro iterativo.** Elijamos ahora el parámetro τ. Nosotros necesitaremos

LEMA 3. *Sean* δ y Δ *las constantes en las desigualdades*

$$\delta\mathscr{D} \leqslant A, \quad R_1\mathscr{D}^{-1}R_2 \leqslant \frac{\Delta}{4}A, \quad \delta > 0. \tag{15}$$

Entonces en la desigualdad (13)

$$\gamma_1 = \delta \Big/ \left(1 + \tau\delta + \tau^2 \frac{\delta\Delta}{4}\right). \tag{16}$$

En efecto, ya que $B^* = \mathscr{D} + \tau R_2$, entonces

$$B\mathscr{D}^{-1}B^* = (\mathscr{D} + \tau R_1)\,\mathscr{D}^{-1}\,(\mathscr{D} + \tau R_2) =$$

$$= \mathscr{D} + \tau\,(R_1 + R_2) + \tau^2 R_1\mathscr{D}^{-1}R_2 = \mathscr{D} + \tau A + \tau^2 R_1\mathscr{D}^{-1}R_2.$$

Utilizando la suposición (15), de aquí obtendremos

$$B\mathscr{D}^{-1}B^* \leqslant (1/\delta + \tau + \tau^2\Delta/4)\,A.$$

El lema está demostrado.

De esta forma, si la información a priori tiene la forma de las constantes δ y Δ en las desigualdades (15), entonces γ_1 se estima por la fórmula (16).

Sustituyendo (16) en (14), obtendremos

$$\| S \|_A^2 \leqslant \varphi(\tau) = 1 - 2\tau\delta \Big/ \left(1 + \tau\delta + \tau^2 \frac{\delta\Delta}{4}\right)^2.$$

Queda por minimizar la función $\varphi(\tau)$. Igualando la derivada $\varphi'(\tau)$, a cero, hallaremos

$$\varphi'(\tau) = \frac{2\delta\left(\tau^2 \frac{\delta\Delta}{4} - 1\right)}{\left(1 + \tau\delta + \tau^2 \frac{\delta\Delta}{4}\right)^2} = 0, \quad \tau_0 = \frac{2}{\sqrt{\delta\Delta}}.$$

Como para $\tau < \tau_0$ la derivada $\varphi'(\tau) < 0$, y para $\tau > \tau_0$ la derivada $\varphi'(\tau) > 0$, entonces para $\tau = \tau_0$ la función $\varphi(\tau)$ alcanza el mínimo, igual a $\varphi(\tau_0) = (1 - \sqrt{\eta})/(1+\sqrt{\eta})$ $\eta = \delta/\Delta$. Así, está demostrado

TEOREMA 6. *Sean A y $\mathscr{D}$ los operadores autoconjugados y definidos positivos en H, y δ, Δ las constantes en* (15). *El método iterativo triangular* (9), (6) *para $\tau = \tau_0 = 2/\sqrt{\delta\Delta}$ converge en H_A, y para el error z_n es válida la estimación $\| z_n \|_A \leqslant \rho^n \| z_0 \|_A$. Para el número n de iteraciones es válida la estimación $n \geqslant n_0(\varepsilon)$,*

$$n_0(\varepsilon) = \ln \varepsilon / \ln \rho,$$

donde $\rho = \left[\frac{1-\sqrt{\eta}}{1+\sqrt{\eta}}\right]^{1/2}$, $\eta = \frac{\delta}{\Delta}$.

4. Estimación de la velocidad de convergencia de los métodos de Seidel y de relajación. El teorema 6 demostrado permite obtener estimaciones de la velocidad de convergencia en H_A de los métodos de Seidel y de relajación superior examinados anteriormente. En el punto 2 del § 1 y el punto 4 del § 2 los métodos indicados fueron aplicados para encontrar una solución aproximada del problema de Dirichlet de diferencias para la ecuación de Poisson sobre la red rectangular

$$\bar{\omega} = \{x_{ij} = (ih_1,\ jh_2),\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,$$
$$h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1,\ 2\}$$
$$\Lambda y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} = -\varphi(x), \quad x \in \omega,$$
$$y(x) = g(x), \quad x \in \gamma.$$

El esquema iterativo de estos métodos tiene la forma (1), donde

$$\mathscr{D}y = \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) \overset{\circ}{y},$$
$$Ly(i,\ j) = -\frac{1}{h_1^2}\overset{\circ}{y}(i-1,\ j) - \frac{1}{h_2^2}\overset{\circ}{y}(i,\ j-1),$$
$$Uy(i,\ j) = -\frac{1}{h_1^2}\overset{\circ}{y}(i+1,\ j) - \frac{1}{h_2^2}\overset{\circ}{y}(i,\ j+1).$$

Para el método de Seidel $\omega = 1$, mientras que para el método de relajación superior ω se ha encontrado por la fórmula (5),

donde δ de la desigualdad (4) fue expresado de la siguiente forma:

$$\delta = \frac{2h_2^2}{h_1^2+h_2^2}\operatorname{sen}^2\frac{\pi h_1}{2l_1} + \frac{2h_1^2}{h_1^2+h_2^2}\operatorname{sen}^2\frac{\pi h_2}{2l_2}. \qquad (17)$$

Reduzcamos el esquema (1) para el ejemplo examinado a la forma (9). Para esto definamos operadores R_1 y R_2:

$$R_1 y = \left(\frac{1}{2}\mathscr{D} + L\right) y = \frac{1}{h_1}\overset{\circ}{y}_{\bar{x}_1} + \frac{1}{h_2}\overset{\circ}{y}_{\bar{x}_2},$$

$$R_2 y = \left(\frac{1}{2}\mathscr{D} + U\right) y = -\frac{1}{h_1}\overset{\circ}{y}_{\bar{x}_1} - \frac{1}{h_2}\overset{\circ}{y}_{\bar{x}_2}.$$

Es evidente que

$$(R_1 + R_2)\, y = Ay = -\Lambda\overset{\circ}{y} = -\overset{\circ}{y}_{\bar{x}_1 x_1} - \overset{\circ}{y}_{\bar{x}_2 x_2}.$$

La conjugación de los operadores R_1 y R_2 uno a otro se establece fácilmente con ayuda de la fórmula de Green de diferencias. Como fue señalado más arriba, al método de Seidel en el esquema (9) le corresponde el valor $\tau = 2$, y al método de relajación superior le corresponde el valor $\tau = 2/\sqrt{\delta(2-\delta)}$, donde δ está definido en (17).

De (11), (14) y del lema 3 se deduce, que para obtener estimaciones de la velocidad de convergencia de estos métodos en H_A se exige hallar δ y Δ de las desigualdades (15). La constante δ ya está hallada. Hallemos Δ. De la definición de los operadores $\mathscr{D}$, R y R_2, obtenemos

$$(R_1\mathscr{D}^{-1}R_2 y,\ y) = 0{,}5\,\frac{h_1^2h_2^2}{h_1^2+h_2^2}(R_2 y,\ R_2 y). \qquad (18)$$

Luego

$$(R_2 y,\ R_2 y) = \frac{1}{h_1^2}(\overset{\circ}{y}_{x_1}^2,\ 1) - \frac{2}{h_1h_2}(\overset{\circ}{y}_{x_1},\ \overset{\circ}{y}_{x_2}) + \frac{1}{h_2^2}(\overset{\circ}{y}_{x_2}^2,\ 1) \leqslant$$

$$\leqslant \left(\frac{1}{h_1^2}+\frac{1}{h_2^2}\right)[(\overset{\circ}{y}_{x_1}^2,\ 1) + (\overset{\circ}{y}_{x_2}^2,\ 1)] \leqslant \frac{h_1^2+h_2^2}{h_1^2h_2^2}(Ay,\ y).$$

Sustituyendo esta estimación en (18), obtendremos

$$(R_1\mathscr{D}^{-1}R_2 y,\ y) \leqslant \frac{1}{2}(Ay,\ y)$$

y, por lo tanto, en la desigualdad (15) $\Delta = 2$.

Estimemos ahora la velocidad de convergencia del método de Seidel y del método de relajación superior.

De (11) obtenemos

$$\| z_n \|_A \leqslant \| S \|_A^n \| z_0 \|_A$$

y, por consiguiente, para alcanzar una exactitud ε es suficiente cumplir $n \geqslant n_0(\varepsilon)$ iteraciones, donde $n_0(\varepsilon) = \ln \varepsilon / \ln \| S \|_A$. De (14) hallaremos

$$n_0(\varepsilon) = 2 \ln \varepsilon / \ln \| S \|_A^2 / \ln \frac{1}{\varepsilon} \Big/ (\tau\gamma_1). \tag{19}$$

Para el método de Seidel de (16) obtendremos ($\tau = 2$)

$$\tau\gamma_1 = 2\delta/(1 + 4\delta) \tag{20}$$

y en el caso particular, cuando $N_1 = N_2 = N$, $l_1 = l_2 = l$, tendremos de (17), (19) y (20)

$$\delta = 2 \operatorname{sen}^2 \frac{\pi}{2N}, \quad \tau\gamma_1 \approx 4 \operatorname{sen}^2 \frac{\pi}{2N} \approx \frac{\pi^2}{N^2},$$

$$n_0(\varepsilon) \approx \frac{N^2}{\pi^2} \ln \frac{1}{\varepsilon} \approx 0{,}1 N^2 \ln \frac{1}{\varepsilon}.$$

En el punto 1 del § 5 del cap. VI para el método explícito de iteración simple, aplicado al caso particular a examinar, fue obtenida la siguiente estimación del número de iteraciones: $n_0(\varepsilon) \approx 0{,}2 N^2 \ln \frac{1}{\varepsilon}$. Comparando estas estimaciones, encontramos, que para el método de Seidel se exige aproximadamente en dos veces menos iteraciones que para el método de iteración simple. El carácter de la dependencia entre el número de iteraciones el de nodos N por una dirección para estos métodos es el mismo—el número de iteraciones es proporcional a N^2.

Examinemos ahora el método de relajación superior. Sustituyendo en (16)

$$\delta = 2 \operatorname{sen}^2 \frac{\pi}{2N}, \quad \Delta = 2 \text{ y } \tau = \frac{2}{\sqrt{\delta(2-\delta)}} = \frac{2}{\operatorname{sen} \frac{\pi}{N}},$$

obtendremos

$$\tau\gamma_1 = \frac{2 \operatorname{tg} \frac{\pi}{2N}}{2 + 2 \operatorname{tg} \frac{\pi}{2N} + \operatorname{tg}^2 \frac{\pi}{2N}} \approx \operatorname{tg} \frac{\pi}{2N} \approx \frac{\pi}{2N}.$$

De (19) hallaremos la siguiente estimación del número de iteraciones para el método de relajación superior:

$$n_0(\varepsilon) \approx \frac{2N}{\pi} \ln \frac{1}{\varepsilon} \approx 0{,}64 N \ln \frac{1}{\varepsilon}, \tag{21}$$

o sea, el número de iteraciones para el método de relajación superior es proporcional al número N de nodos por una dirección.

Como conclusión mostremos una estimación para el número de iteraciones, el cual se deduce del teorema 6. Siendo el valor del parámetro τ

$$\tau = \tau_0 = \frac{2}{\sqrt{\delta\Delta}} = \frac{1}{\operatorname{sen} \frac{\pi}{2N}}$$

para el número de iteraciones será válida la estimación

$$n \geqslant n_0(\varepsilon) = \ln \frac{1}{\varepsilon} \Big/ \sin \frac{\pi}{2N} \approx 0{,}6N \ln \frac{1}{\varepsilon}.$$

Notemos, que la estimación (21) está algo aumentada. Para convencerse de esto, es necesario comparar los valores de $n_0(\varepsilon)$, calculados por la fórmula (21), con el número de iteraciones, citado en el punto 4 del § 2. Esto está relacionado con que aquí el número de iteraciones se estimó de la desigualdad $\| S \|_A^n \leqslant \varepsilon$, y en el punto 4 del § 2 las iteraciones se realizaron hasta el cumplimiento de la condición $\| S^n \|_A \leqslant \varepsilon$.

Capítulo X

Método alternativo triangular

En este capítulo se estudia el método iterativo alternativo triangular*) de resolución de una ecuación operacional con un operador autoconjugado. En el § 1 se expone la teoría general del método, está descrita la contrucción del esquema iterativo e indicado el conjunto de los parámetros iterativos. Se ilustra el método con el ejemplo del problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo. En el § 2 este método se aplica a la resolución de ecuaciones elípticas en diferencias con coeficientes variables y con derivadas mixtas en un rectángulo. Para resolver una ecuación elíptica con coeficientes variables sobre una red no uniforme en una región arbitraria en el § 3 está construida una variante del método alternativo triangular.

§ 1. Teoría general del método

1. Esquema iterativo. En el § 3 del cap. IX fue estudiado el método iterativo triangular de resolución de la ecuación

$$Au = f. \tag{1}$$

El esquema iterativo de este método tiene la forma

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, y_0 \in H, \tag{2}$$

donde $\tau_k \equiv \tau$, y el operador $B = B_1 = \mathscr{D} + \tau R_1$ se define por el siguiente desarrollo del operador A en la suma de los operadores:

$$A = R_1 + R_2, \quad R_1 = R_2^*, \quad A = A^* > 0. \tag{3}$$

*) El método fue propuesto por A. A. Samarski en el año 1964 (véase R de MC y FM 4, No. 3, 1964) y perfeccionado en [8].

Respecto al operador $\mathscr{D}$ se supone, que él es autoconjugado y definido positivo en H, es decir,

$$\mathscr{D} = \mathscr{D}^* > 0. \tag{4}$$

El método iterativo triangular corresponde a la clase de los métodos, cuyos parámetros iterativos se eligen teniendo en cuenta la información a priori sobre los operadores del esquema iterativo. Para el método triangular la información primaria consiste en prefijar las constantes δ y Δ de las desigualdades

$$\delta\mathscr{D} \leqslant A, \quad R_1\mathscr{D}^{-1}R_2 \leqslant \frac{\Delta}{4} A, \quad \delta > 0. \tag{5}$$

El parámetro τ hallado en el punto 3 del § 3 del cap. IX, permite obtener una exactitud ε después de $n_0 = O \times \times (\ln (1/\varepsilon) \sqrt{\eta})$ iteraciones, donde $\eta = \delta/\Delta$.

Señalemos que la no autoconjugación del operador B no permite utilizar el conjunto de los parámetros τ_k en el esquema iterativo (2) y con ello aumentar la velocidad de convergencia del método. Sin embargo, la simpleza de la construcción del operador B y la posibilidad del desarrollo (3) para los operadores A de cualquier estructura estimularon el estudio de posibles modificaciones del método triangular. Como resultado fue construido el método alternativo triangular, el cual combina la universalidad de la construcción del operador B con la posibilidad de elegir el conjunto de los parámetros τ_k en el esquema (2).

Abordemos el estudio del método alternativo triangular. El esquema iterativo del método tiene la forma (2), donde el operador B se define de la forma siguiente:

$$B = (\mathscr{D} + \omega R_1)\,\mathscr{D}^{-1}\,(\mathscr{D} + \omega R_2), \quad \omega > 0. \tag{6}$$

Aquí ω es el parámetro iterativo, sujeto a definición. A continuación vamos a suponer que para el esquema (2), (6) son cumplidas las condiciones (3), (4) y son dadas δ y Δ en las desigualdades (5).

Señalemos algunas propiedades del operador B, definido por la relación (6). Si al operador $\mathscr{D} + \omega R_1$ le corresponde una matriz triangular, y $\mathscr{D}$ es diagonal, entonces a B le corresponde el producto de dos matrices triangulares y una diagonal. En este caso la inversión del operador B no es un problema complejo.

Mostremos que el operador B es autoconjugado en H, y si el operador $\mathscr{D}$ es acotado, entonces B es definido positivo. Realmente, en virtud de (3) tenemos las igualdades

$$(Au,\ u) = 2\,(R_1u,\ u) = (2\,(R_2u,\ u) > 0.$$

De aquí y de (4) se deduce que los operadores $B_1 = \mathscr{D} + \omega R_1$ y $B_2 = \mathscr{D} + \omega R_2$ son conjugados y definidos positivos: $B_1^* = (\mathscr{D} + \omega R_1)^* = \mathscr{D} + \omega R_2 = B_2$, $B_\alpha > \mathscr{D} > 0$, $\alpha = 1, 2$ por eso

$$B^* = (B_1\mathscr{D}^{-1}B_2)^* = B_2^*\mathscr{D}^{-1}B_1^* = B_1\mathscr{D}^{-1}B_2 = B.$$

Seguidamente, como $\mathscr{D}$ es un operador autoconjugado acotado y definido positivo, entonces el operador inverso $\mathscr{D}^{-1}$ será definido positivo en H. Por consiguiente, utilizando la desigualdad $(\mathscr{D}^{-1}x,\ x) \geqslant d\,(x,\ x)$, $d > 0$, la cual expresa la definición positiva del operador $\mathscr{D}^{-1}$, tenemos

$$(Bu,\ u) = (\mathscr{D}^{-1}B_2u,\ B_2u) \geqslant d\,\|\,B_2u\,\|^2 > 0.$$

De (2) y (6) se ve, que para determinar y_{k+1} siendo dado y_k es necesario resolver la ecuación

$$(\mathscr{D} + \omega R_1)\,\mathscr{D}^{-1}\,(\mathscr{D} + \omega R_2)\,y_{k+1} = \varphi_k,\quad k = 0,\ 1,\ \ldots,$$

donde $\varphi_k = By_k - \tau_{k+1}\,(Ay_k - f)$. Esta operación se reduce a la resolución de dos ecuaciones

$$(\mathscr{D} + \omega R_1)\,v = \varphi_k,\quad (\mathscr{D} + \omega R_2)\,y_{k+1} = \mathscr{D}v.$$

Es posible el segundo algoritmo de realización del esquema (2), (6) basado en su escritura en forma de un esquema con corrección

$$y_{k+1} = y_k - \tau_{k+1}w_k,\quad Bw_k = r_k,$$

donde $r_k = Ay_k - f$ es el defecto. La corrección w_k se encuentra mediante la resolución de las dos ecuaciones

$$(\mathscr{D} + \omega R_1)\,\bar{\omega}_k = r_k,\quad (\mathscr{D} + \omega R_2)\,w_k = \mathscr{D}\bar{w}_k.$$

En este algoritmo nosotros nos hemos liberado de la necesidad de calcular By_k, pero estamos obligados a conservar simultáneamente y_k y las magnitudes intermedias r_k, $\bar{\omega}_k$ y ω_k.

2. Elección de los parámetros iterativos. Ocupémonos ahora en investigar la convergencia del esquema iterativo (2), (6). Puesto que los operadores A y B son autoconjugados y definidos positivos en H, entonces se puede estudiar la convergencia en H_D, donde en calidad de D se toma uno de los operadores A, B o $AB^{-1}A$ (en el último caso B debe

ser un operador acotado). Para el operador indicado D, el operador $DB^{-1}A$, evidentemente, será autoconjugado en H y, por consiguiente, de acuerdo con la clasificación del capítulo VI, tenemos el esquema iterativo para el caso autoconjugado.

Utilizando los resultados del § 2 del cap. VI, nosotros inmediatamente podemos indicar para el esquma (2), (6) el conjunto óptimo de los parámetros iterativos τ_k. Sean γ_1 y γ_2 tomados de las desigualdades

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0. \tag{7}$$

Entonces el conjunto de los parámetros de Chebishev $\{\tau_k\}$ se define mediante las fórmulas

$$\tau_k = \frac{\tau_0}{1+\rho_0\mu_k}, \quad \mu_k \in \mathfrak{M}_n^* = \left\{\cos\frac{(2i-1)\pi}{2n}, \ 1 \leqslant i \leqslant n\right\}, \quad 1 \leqslant k \leqslant n, \tag{8}$$

$$\tau_0 = \frac{2}{\gamma_1+\gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad n \geqslant n_0(\varepsilon) = \frac{\ln(0{,}5\varepsilon)}{\ln\rho_1}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

y para el error $z_n = y_n - u$ del método iterativo (2), (6), (8) tiene lugar la estimación

$$\|z_n\|_D \leqslant q_n \|z_0\|_D, \quad q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}} \leqslant \varepsilon, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}. \tag{9}$$

Este es un resultado de la teoría general de los métodos iterativos de dos capas. Para el esquema (2), (6) la información a priori son las constantes δ y Δ en las desigualdades (5). Por eso uno de los problemas es la obtención de expresiones para γ_1 y γ_2 mediante δ y Δ. Después, por cuanto el operador B depende del parámetro iterativo ω, entonces γ_1 y γ_2 son funciones de ω: $\gamma_1 = \gamma_1(\omega)$, $\gamma_2 = \gamma_2(\omega)$. Ya que de la estimación (9) se deduce que la velocidad máxima de convergencia será cuando la relación $\xi = \gamma_1/\gamma_2$ sea máxima, entonces llegamos al problema sobre la elección del parámetro ω de la condición de máximo de ξ. El problema formulado lo resuelve el

LEMA 1. *Sean cumplidas las condiciones* (3), (4), *el operador* B *definido por la fórmula* (6) *y dadas las constantes*

δ y Δ *en las desigualdades* (5). *Entonces en las desigualdades* (7) *tenemos*

$$\gamma_1 = \delta \Big/ \left(1 + \omega\delta + \tfrac{1}{4}\,\omega^2\delta\Delta\right), \quad \gamma_2 = 1/(2\omega). \tag{10}$$

La relación $\xi = \gamma_1/\gamma_2$ *es máximo, si*

$$\omega = \omega_0 = 2/\sqrt{\delta\Delta}, \tag{11}$$

con esto

$$\gamma_1 = \frac{\delta}{2(1+\sqrt{\eta})}, \quad \gamma_2 = \frac{\delta}{4\sqrt{\eta}}, \quad \xi = \frac{2\sqrt{\eta}}{1+\sqrt{\eta}}, \quad \eta = \frac{\delta}{\Delta}. \tag{12}$$

En efecto, escribamos el operador B en la forma $B = (\mathscr{D} + \omega R_1)\,\mathscr{D}^{-1}(\mathscr{D} + \omega R_2) = \mathscr{D} + \omega(R_1 + R_2) + \omega^2 R_1 \mathscr{D}^{-1} R_2$. (13)

Teniendo en cuenta, que $A = R_1 + R_2 \geqslant \delta\mathscr{D}$ ó $\mathscr{D} \leqslant \frac{1}{\delta} A$ obtenemos para B la estimación por encima

$$B \leqslant \left(\frac{1}{\delta} + \omega + \frac{1}{4}\omega^2\Delta\right) A = \frac{1}{\gamma_1} A,$$

es decir $A/\gamma_1 B$, donde γ_1 está definido en (10).

Transformemos ahora la fórmula (13):

$$B = \mathscr{D} - \omega(R_1 + R_2) + \omega^2 R_1 \mathscr{D}^{-1} R_2 + 2\omega(R_1 + R_2) = (\mathscr{D} - \omega R_1)\,\mathscr{D}^{-1}(\mathscr{D} - \omega R_2) + 2\omega A.$$

De aquí se deduce

$$(By,\ y) = 2\omega(Ay,\ y) + (\mathscr{D}^{-1}(\mathscr{D} - \omega R_2)\,y,\ (\mathscr{D} - \omega R_2)y).$$

Utilizando la definición positiva del operador $\mathscr{D}^{-1}$, obtenemos $(By,\ y) \geqslant 2\omega(Ay,\ y)$ es decir, $A \leqslant \gamma_2 B$. Así, están hallados γ_1 y γ_2. Examinemos a continuación la razón

$$\xi = \xi(\omega) = \gamma_1/\gamma_2 = 2\omega\delta \Big/ \left(1 + \omega\delta + \frac{\omega^2\delta\Delta}{4}\right).$$

Igualando a cero la derivada

$$\xi'(\omega) = \frac{2\delta(1 - \omega^2\delta\Delta/4)}{\left(1 + \omega\delta\,\frac{\omega^2\delta\Delta}{4}\right)^2},$$

encontramos $\omega = \omega_0 = 2/\sqrt{\delta\Delta}$. En este punto se alcanza el máximo de $\xi(\omega)$, puesto que $\xi''(\omega_0) < 0$. Sustituyendo

el ω hallado en (10), obtenemos (12). Mostremos que $\delta \leqslant \Delta$ y $\eta \leqslant 1$. En efecto, utilizando la igualdad $(Ax, x) = -2(R_2 x, x)$ y la desigualdad de Cauchy—Buniakovski, de (5) obtenemos

$$\delta(\mathscr{D}x, x) \leqslant (Ax, x) = \frac{(Ax, x)^2}{(Ax, x)} = 4\frac{(R_2 x, x)^2}{(Ax, x)} =$$

$$= 4\frac{(\mathscr{D}^{-1/2}R_2 x, \mathscr{D}^{1/2}x)^2}{(Ax, x)} \leqslant 4\frac{(\mathscr{D}^{-1/2}R_2 x, \mathscr{D}^{-1/2}R_2 x)}{(Ax, x)} \times$$

$$\times (\mathscr{D}^{1/2}x, \mathscr{D}^{1/2}x) = 4\frac{(R_1\mathscr{D}^{-1}R_2 x, x)}{(Ax, x)}(\mathscr{D}x, x) \leqslant \Delta(\mathscr{D}x, x),$$

lo que se exigía demostrar. El lema está demostrado.

TEOREMA 1. *Sean cumplidas las condiciones del lema* 1. *Entonces para el método alternativo triangular* (2), (6), (11) *de los parámetros de Chébishev* τ_k, *definidos por las fórmulas* (8) *y* (12), *es válida la estimación* (9). *Para el cumplimiento de la desigualdad* $\| z_n \|_D \leqslant \varepsilon \| z_0 \|_D$ *son suficientes n iteraciones, donde* $n \geqslant n_0(\varepsilon)$, $n_0(\varepsilon) = \ln\frac{2}{\varepsilon} \Big/ (2\sqrt{2}\sqrt[4]{\eta})$, $\eta = \delta/\Delta$. *Aquí* $D = A$, B *o* $AB^{-1}A$.

Para la demostración del teorema se debe utilizar el lema 1 y las fórmulas (8) para los parámetros iterativos y el número de iteraciones.

Examinemos ahora un procedimiento, el cual se utiliza durante la construcción de los esquemas iterativos implícitos. Sea dado un operador R autoconjugado y definido positivo en H, energéticamente equivalente al operador A con constantes c_1 y c_2:

$$c_1 R \leqslant A \leqslant c_2 R, \quad c_1 > 0, \tag{14}$$

y al operador B con constantes $\mathring{\gamma}_1$ y $\mathring{\gamma}_2$:

$$\mathring{\gamma}_1 B \leqslant R \leqslant \mathring{\gamma}_2 B, \quad \mathring{\gamma}_1 > 0. \tag{15}$$

Supongamos que los operadores A y B son autoconjugados. De (14) y (15) para ellos obtendremos las desigualdades siguientes: $\gamma_1 B \leqslant A \leqslant \gamma_2 B$, $\gamma_1 = c_1\mathring{\gamma}_1$, $\gamma_2 = c_2\mathring{\gamma}_2$. El procedimiento expuesto permite al construir el operador B partir no del desarrollo (3) del operador A, sino del desarrollo del operador R, el cual puede ser elegido un mismo para una gran clase de operadores A diferentes. Con esto las constantes $\mathring{\gamma}_1$ y $\mathring{\gamma}_2$ en (15) pueden ser halladas una vez, y todo

el problema de la obtención de información a priori para el método se reduce a encontrar c_1 y c_2 en (14).

Así pues, sea el operador R representado en forma de la suma de los operadores conjugados R_1 y R_2:

$$R = R^* > 0, \quad R = R_1 + R_2, \quad R_1 = R_2^*, \tag{16}$$

y en lugar de (5) tienen lugar las desigualdades

$$\delta \mathscr{D} \leqslant R, \quad R_1 \mathscr{D}^{-1} R_2 \leqslant \frac{\Delta}{4} R, \quad \delta > 0. \tag{17}$$

Construyamos el operador B para el esquema (2) por la fórmula (6). Entonces en virtud del lema 1 para $\omega = \omega_0 = 2/\sqrt{\delta\Delta}$ en las desigualdades (15) tenemos

$$\overset{\circ}{\gamma}_1 = \frac{\delta}{2(1+\sqrt{\eta})}, \quad \overset{\circ}{\gamma}_2 = \frac{\delta}{4\sqrt{\eta}}, \quad \overset{\circ}{\xi} = \frac{\overset{\circ}{\gamma}_1}{\overset{\circ}{\gamma}_2} = \frac{2\sqrt{\eta}}{1+\sqrt{\eta}}, \quad \eta = \frac{\delta}{\Delta}. \tag{18}$$

De aquí se deduce

TEOREMA 2. *Sean* $A = A^* > 0$, $\mathscr{D} = \mathscr{D}^* > 0$, *cumplidas las condiciones* (16) *y prefijadas* c_1 *y* c_2 *en* (14) *y* δ, Δ *en* (17). *Entonces para el método alternativo triangular* (2), (6), (8), (11) *con los parámetros de Chebishev* τ_k, *donde* $\gamma_1 = c_1\overset{\circ}{\gamma}_1$ *y* $\gamma_2 = c_2\overset{\circ}{\gamma}_2$ *y* $\overset{\circ}{\gamma}_1$, $\overset{\circ}{\gamma}_2$ *definidos en* (18), *es válida la estimación* (9). *Para cumplir la desigualdad* $\| z_n \|_D \leqslant \leqslant \varepsilon \| z_0 \|_D$ *son suficientes* n *iteraciones, donde*

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \frac{\ln(2/\varepsilon)}{2\sqrt{2}\sqrt[4]{\eta}} \sqrt{\frac{c_2}{c_1}}, \quad \eta = \frac{\delta}{\Delta}.$$

3. Método para encontrar las magnitudes iniciales δ y Δ. De los teoremas 1 y 2 se deduce que para aplicar el método alternativo triangular se exige prefijar dos números δ y Δ en las desigualdades (5) o (17). En los ejemplos examinados más abajo de ecuaciones elípticas reticulares estas constantes serán halladas en forma explícita o serán indicados los algoritmos para su cálculo. En esto, naturalmente, se utiliza la estructura de los operadores A, R_1, R_2 y $\mathscr{D}$. Para la teoría general de los métodos iterativos, la cual no toma en consideración la estructura concreta de los operadores, se necesita proponer un procedimiento general para obtener la información a priori, requerida para realizar dicho método.

Este procedimiento puede estar basado en la utilización de la propiedad asintótica de los métodos iterativos de tipo variacional (véase el punto 5 del § 1 del cap. VIII). Sean los operadores A y B autoconjugados y definidos positivos en H. Si en el esquema iterativo

$$B\frac{v_{k+1}-v_k}{\tau_{k+1}}+Av_k=0,\quad k=0,\ 1,\ \ldots,\ v_0\neq 0 \tag{19}$$

se eligen los parámetros τ_{k+1} por la fórmula del método del descenso más rápido

$$\tau_{k+1}=\frac{(w_k,\ r_k)}{(Aw_k,\ w_k)},\quad k=0,\ 1,\ \ldots,\ r_k=Av_k,\ Bw_k=r_k, \tag{20}$$

y para un número suficientemente grande de iteraciones n hallamos las raíces $x_1\leqslant x_2$ de la ecuación

$$(1-\tau_n x)(1-\tau_{n-1}x)=\rho_n\rho_{n-1},\ \rho_n=\frac{\|v_n\|_A}{\|v_{n-1}\|_A}, \tag{21}$$

donde $\|\cdot\|_A$ es la norma en H_A, entonces x_1 y x_2 serán las aproximaciones a γ_1 y γ_2 en las desigualdades (7) respectivamente por encima y por debajo.

Utilicemos el procedimiento descrito. Examinemos el esquema iterativo (2), (3), (6). Notemos que en virtud del lema 1 para la desigualdad (7) $\gamma_2=1/(2\omega)$, y sólo γ_1 depende de los datos apriorísticos. Nosotros trataremos, sin hallar por separado δ y Δ para las desigualdades (5), encontrar inmediatamente la expresión para γ_1 como función del parámetro iterativo ω. A esta expresión nosotros le daremos la forma (10), indicando los valores correspondientes de δ y Δ. Entonces del lema 1 hallaremos ω_0 mediante la fórmula (11), y γ_1, γ_2 correspondientes a ω_0, mediante la fórmula (12). Para el conjunto de los parámetros τ_k utilizaremos (8).

Encontremos la expresión buscada para γ_1. Tomemos $\omega=0$ y aplicando el método (19) — (21) hallaremos x_1. Considerando que en (19), (20) se ha realizado una cantidad suficiente de iteraciones, y teniendo en cuenta que para $\omega=0$ el operador $B=\mathscr{D}$, obtendremos la desigualdad aproximada

$$x_1\mathscr{D}\leqslant A,\quad x_1>0. \tag{22}$$

A continuación, tomemos $\omega=\omega_1>0$ y mediante el método (19) — (21) hallaremos $\bar{x}_1$, así que $\bar{x}_1>0$ y

$$\bar{x}_1 B\leqslant A\ \text{ o }\ \bar{x}_1(\mathscr{D}+\omega_1 A+\omega_1^2 R_1\mathscr{D}^{-1}R_2)\leqslant A, \tag{23}$$

al mismo tiempo se ve, que $\bar{x}_1\omega_1 < 1$. Escribamos (23) en la forma

$$\bar{x}_1\mathscr{D} + \bar{x}_1\omega_1^2 R_1\mathscr{D}^{-1}R_2 \leqslant (1 - \bar{x}_1\omega_1)\,A$$

y sumémosla con (22), la cual multiplicamos con anticipación por cierto coeficiente $\alpha > 0$ indeterminado por ahora. Obtendremos

$$(\alpha x_1 + \bar{x}_1)\,\mathscr{D} + \bar{x}_1\omega_1^2 R_1\mathscr{D}^{-1}R_2 \leqslant (1 - \bar{x}_1\omega_1 + \alpha)\,A. \qquad (24)$$

Dividamos esta desigualdad por $\alpha x_1 + \bar{x}_1$, añadamos a los miembros derecho e izquierdo el sumando ωA y elijamos α de la condición

$$\bar{x}_1\omega_1^2 = \omega^2\,(\alpha x_1 + \bar{x}_1); \qquad (25)$$

entonces la desigualdad transformada tendrá la forma

$$\mathscr{D} + \omega A + \omega^2 R_1\mathscr{D}^{-1}R_2 = B \leqslant \frac{1}{\gamma_1}\,A,$$

donde

$$\frac{1}{\gamma_1} = \frac{1}{\gamma_1(\omega)} = \omega + \frac{1 - \bar{x}_1\omega_1 + \alpha}{\alpha x_1 + \bar{x}_1}\,. \qquad (26)$$

De (25) hallaremos α:

$$\alpha = x_1\,(\omega_1^2 - \omega^2)/(\omega^2 x_1).$$

Ya que α debe ser positivo, entonces la expresión (26) tendrá lugar para $0 < \omega < \omega_1$. Sustituyendo el α hallado en (26), obtendremos

$$\frac{1}{\gamma_1} = \frac{1}{x_1} + \omega + \frac{x_1 - \bar{x}_1 - x_1\bar{x}_1\omega_1}{x_1\bar{x}_1\omega_1^2}\,\omega^2.$$

Comparando esta expresión con (10), obtenemos, que en calidad de δ y Δ se pueden tomar

$$\delta = x_1, \quad \Delta = 4\,\frac{x_1 - \bar{x}_1 - x_1\bar{x}_1\omega_1}{x_1\bar{x}_1\omega_1^2}\,.$$

Señalemos que ω_0, hallado por los δ y Δ indicados de acuerdo con (11), pertenecerá al intervalo $(0, \omega_1)$, si se cumple la desigualdad $2\bar{x}_1 \leqslant x_1\,(1 - \bar{x}_1\omega_1)$. Si esta desigualdad no se cumple, entonces se debe aumentar ω_1 y realizar los cálculos indicados de nuevo (se recomienda tomar $\omega_1 = 2/x_1$).

4. **Problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo.** Ilustremos el método alternativo triangular en el ejemplo del problema de Dirichlet de diferencias para la ecuación de Poisson en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$:

$$\Lambda y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} = -\varphi(x), \quad x \in \omega,$$
$$y(x) = g(x), \quad x \in \gamma \tag{27}$$

sobre la red $\bar{\omega} = \{x_{ij} = (ih_1, jh_2) \in \bar{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1, 2\}$ con frontera γ.

Para el ejemplo dado H es el espacio de las funciones reticulares, definidas sobre ω, con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2.$$

El operador A se define por la igualdad $Ay = -\Lambda \overset{\circ}{y}$, donde $y \in H$, $\overset{\circ}{y} \in \overset{\circ}{H}$, e $y(x) = \overset{\circ}{y}(x)$, $x \in \omega$, e $\overset{\circ}{y}(x) = 0$ para $x \in \gamma$. El segundo miembro f lo definiremos de la forma usual: $f(x) = \varphi(x) + \frac{1}{h_1^2}\varphi_1(x) + \frac{1}{h_2^2}\varphi_2(x)$, donde

$$\varphi_1(x) = \begin{cases} g(0, x_2), & x_1 = h_1, \\ 0, & 2h \leqslant x_1 \leqslant l_1 - 2h_1, \\ g(l_1, x_2), & x_1 = l_1 - h_1, \end{cases}$$

$$\varphi_2(x) = \begin{cases} g(x_1, 0), & x_2 = h_2, \\ 0, & 2h_2 \leqslant x_2 \leqslant l_2 - 2h_2, \\ g(x_1, l_2), & x_2 = l_2 - h_2. \end{cases}$$

Entonces el problema (27) se escribe en la forma de la ecuación (1).

El operador A es autoconjugado y definido positivo en H, puesto que él corresponde al operador de Laplace de diferencias para las condiciones de contorno de Dirichlet.

Ocupémonos ahora de la construcción del operador B. Vamos a examinar la variante clásica del método alternativo triangular, para la cual en (6) pondremos

$$\mathscr{D} = E. \tag{28}$$

Definamos ahora los operadores de diferencias $\mathscr{R}_1$ y $\mathscr{R}_2$, los cuales actúan sobre las funciones reticulares definidas sobre $\bar{\omega}$, de la siguiente forma:

$$\mathscr{R}_1 y = -\sum_{\alpha=1}^{2} \frac{1}{h_\alpha} y_{\bar{x}_\alpha}, \quad \mathscr{R}_2 y = \sum_{\alpha=1}^{2} \frac{1}{h_\alpha} y_{x_\alpha}, \quad x \in \omega.$$

Es evidente, que $\mathscr{R}_1 + \mathscr{R}_2 = A$. Utilizando las fórmulas de Green de diferencias, obtendremos fácilmente, que para las funciones reticulares $\overset{\circ}{y}(x) \in \overset{\circ}{H}$, $\overset{\circ}{u}(x) \in \overset{\circ}{H}$, es decir, definidas sobre $\bar{\omega}$ y que se anulan sobre γ, tiene lugar la igualdad

$$(\mathscr{R}_1 \overset{\circ}{y}, \overset{\circ}{u}) = (\overset{\circ}{y}, \mathscr{R}_2 \overset{\circ}{u}). \tag{29}$$

Definamos sobre H los operadores $\mathscr{R}_1$ y $\mathscr{R}_2$ de la siguiente forma: $\mathscr{R}_\alpha y = -\mathscr{R}_\alpha \overset{\circ}{y}$, $\alpha = 1, 2$, donde $y \in H$, $\overset{\circ}{y} \in \overset{\circ}{H}$ e $y(x) = \overset{\circ}{y}(x)$, $x \in \omega$. Entonces en virtud de la definición de operadores de diferencias $\mathscr{R}_\alpha$ y de la igualdad (29) se cumplen las condiciones (3), es decir, $A = R_1 + R_2$, $R_1 = R_2^*$. Teniendo en cuenta (28), de (6) obtendremos la siguiente forma del operador B:

$$B = (E + \omega R_1)(E + \omega R_2).$$

Hallemos ahora la información a priori necesaria para realizar el método alternativo triangular. En el caso dado ella posee la forma de las constantes δ y Δ en las desigualdades $\delta E \leqslant A$ y $R_1 R_2 \leqslant (\Delta/4) A$. Es evidente que en calidad de δ se puede tomar el valor propio mínimo del operador de Laplace de diferencias

$$\delta = \frac{4}{h_1^2} \operatorname{sen}^2 \frac{\pi h_1}{2l_1} + \frac{4}{h_2^2} \operatorname{sen}^2 \frac{\omega h_2}{2l_2}.$$

La estimación para Δ está hallada en el punto 4 del § 3 del cap. IX (los operadores R_1 y R_2, definidos aquí y allá, coinciden). Tenemos $\Delta = 4/h_1^2 + 4/h_2^2$.

Así pues, está obtenida la información necesaria sobre δ y Δ. Del lema 1 encontramos el valor óptimo para el parámetro ω_0, y también γ_1 y γ_2. Los parámetros itera-

tivos τ_k se calculan por las fórmulas (8). El el caso particular, cuando $N_1 = N_2 = N$, $l_1 = l_2 = l$, obtenemos

$$\delta = \frac{8}{h^2}\operatorname{sen}^2\frac{\pi}{2N}, \quad \Delta = \frac{8}{h^2}, \quad \eta = \frac{\delta}{\Delta} = \operatorname{sen}^2\frac{\pi}{2N},$$

$$\xi = \frac{2\sqrt{\eta}}{1+\sqrt{\eta}} \approx 2\sqrt{\eta} = 2\operatorname{sen}\frac{\pi}{2N} \approx \frac{\pi}{N}, \quad \omega_0 = \frac{h^2}{4\operatorname{sen}\frac{\pi h}{2l}}.$$

En este caso del teorema 1 para el número de iteraciones n tendremos $n \geqslant n_0(\varepsilon)$, donde

$$n_0(\varepsilon) = \frac{\ln(2/\varepsilon)}{2\sqrt{2}\sqrt[4]{\eta}} \approx \frac{\sqrt{N}}{2\sqrt{\pi}}\ln\frac{2}{\varepsilon} \approx 0{,}28\sqrt{N}\ln\frac{2}{\varepsilon},$$

es decir, el número de iteraciones es proporcional a la raíz de cuarto grado del número de incógnitas en el problema.

En el punto 4 del § 3 del cap. IX para el método de relajación superior, aplicado a la resolución del problema de diferencias (27), fue obtenida la siguiente estimación del número de iteraciones:

$$n \geqslant n_0(\varepsilon) \approx 0{,}64N\ln(1/\varepsilon), \quad (N = l/h).$$

La comparación del método de relajación con el método alternativo triangular muestra una superioridad explícita del último. Aunque en la realización de un paso iterativo en el método alternativo triangular es necesario gastar dos veces más operaciones aritméticas que en el método de relajación, el primero posee una ventaja esencial en el número de iteraciones, lo cual asegura la efectividad general de este método.

Ponemos ahora el número de iteraciones para el método alternativo triangular con los parámetros de Chebishev examinado aquí para el problema de diferencias (27), en dependencia del número de nodos N por una dirección de la red cuadrada $\overline{\omega}$ para $\varepsilon = 10^{-4}$:

$N = 32 \qquad n = 16$

$N = 64 \qquad n = 23$

$N = 128 \qquad n = 32$

La comparación con el número de iteraciones del método de relajación superior, la cual está expuesta en el punto 4

del § 2 del cap. IX, muestra que el método de relajación exige aproximadamente en 3,5—7,5 veces más iteraciones que el método alternativo triangular.

OBSERVACION 1. Si en lugar del rectángulo $\bar{G}$ examinamos el paralelepípedo p-dimensional $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2, \ldots, p\}$, y en él el problema de Dirichlet de diferencias para la ecuación de Poisson

$$\Lambda y = \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha x_\alpha} = -\varphi(x), \quad x \in \omega,$$

$$y(x) = g(x), \quad x \in \gamma,$$

sobre la red rectangular $\bar{\omega} = \{x_i = (i_1 h_1, i_2 h_2, \ldots, i_p h_p) \in \bar{G},\ 0 \leqslant i_\alpha \leqslant N_\alpha,\ h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2, \ldots, p\}$, entonces los operadores de diferencias $\mathscr{R}_1$ y $\mathscr{R}_2$ se definen de la forma siguiente:

$$\mathscr{R}_1 y = -\sum_{\alpha=1}^{p} \frac{1}{h_\alpha} y_{\bar{x}_\alpha}, \quad \mathscr{R}_2 y = \sum_{\alpha=1}^{p} \frac{1}{h_\alpha} y_{x_\alpha}, \quad x \in \omega.$$

En este caso en las desigualdades (5) para $\mathscr{D} = E$ se debe poner

$$\delta = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2 l_\alpha}, \quad \Delta = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2},$$

ya que

$$\| \mathring{\mathscr{R}}_2 \mathring{y} \|^2 = (\mathscr{R}_1 \mathscr{R}_2 \mathring{y}, \mathring{y}) = \Big\| \sum_{\alpha=1}^{p} \frac{1}{h_\alpha} \mathring{y}_{x_\alpha} \Big\| \leqslant$$

$$\leqslant \sum_{\alpha=1}^{p} \frac{1}{h_\alpha^2} \sum_{\alpha=1}^{p} \| \mathring{y}_{x_\alpha} \|^2 \leqslant \Big(\sum_{\alpha=1}^{p} \frac{1}{h_\alpha^2} \Big) (A\mathring{y}, \mathring{y}).$$

Con esto cuando $N_1 = N_2 = \ldots = N_p = N$, $l_1 = l_2 = \ldots = l_p = l$ para el número de iteraciones tenemos la estimación

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) \approx \frac{\sqrt{N}}{2\sqrt{\pi}} \ln \frac{2}{\varepsilon} \approx 0{,}28 \sqrt{N} \ln \frac{2}{\varepsilon},$$

la cual no depende del número de dimensiones p.

Examinemos ahora las preguntas relacionadas con la realización del método alternativo triangular para el pro-

blema (27). En el punto 1 fueron expuestos dos algoritmos para encontrar y_{k+1} por y_k prefijado para el esquema iterativo del método. Examinemos primeramente el segundo algoritmo. Para $\mathscr{D} = E$ él posee la forma siguiente:

$$r_k = Ay_k - f,$$
$$(E + \omega_0 R_1)\,\bar{w}_k = r_k, \quad (E + \omega_0 R_2)\,\omega_k = \bar{w}_k, \qquad (30)$$
$$y_{k+1} = y_k - \tau_{k+1} w_k, \quad k = 0, 1, \ldots$$

Este algoritmo se debe emplear cuando los parámetros τ_k no se eligen por las fórmulas (8), sino por las fórmulas de los métodos iterativos de tipo variacional.

Utilizando la definición de los operadores A, R_1 y R_2 por medio de los operadores de diferencias Λ, $\mathscr{R}_1$ y $\mathscr{R}_2$, se pueden escribir las fórmulas (30) en la forma siguiente:

$$r_k(i, j) = \left(\frac{2}{h_1^2} + \frac{2}{h_2^2}\right) y_k(i, j) - \frac{1}{h_1^2}[y_k(i-1, j) + y_k(i+1, j)] - \frac{1}{h_2^2}[y_k(i, j-1) + y_k(i, j+1)] - \varphi(i, j), \qquad (31)$$

$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1, \quad y_k|_\gamma = g.$$

$$\bar{w}_k(i, j) = \alpha \bar{w}_k(i-1, j) + \beta \bar{w}_k(i, j-1) + \varkappa r_k(i, j),$$
$$i = 1, 2, \ldots, N_1 - 1, \quad j = 1, 2, \ldots, N_2 - 1, \qquad (32)$$
$$w_k(0, j) = 0, \quad 1 \leqslant j \leqslant N_2 - 1, \quad \bar{\omega}_k(i, 0) = 0, \quad 1 \leqslant i \leqslant N_1 - 1.$$

Aquí el cálculo se lleva a efecto, comenzando por el punto $i = 1$, $j = 1$ bien sea por las filas de la red $\bar{\omega}$, es decir, al crecer i para j fijo, o bien por las columnas al crecer j para i fijo.

$$w_k(i, j) = \alpha w_k(i+1, j) + \beta w_k(i, j+1) + \varkappa \bar{w}_k(i, j),$$
$$i = N_1 - 1, N_1 - 2, \ldots, 1, \quad j = N_2 - 1, N_2 - 2, \ldots, 1, \qquad (33)$$
$$w_k(N_1, j) = 0, \quad 1 \leqslant j \leqslant N_2 - 1, \quad w_k(i, N_2) = 0, \quad 1 \leqslant i \leqslant N_1 - 1.$$

Aquí el cálculo se lleva a efecto, comenzando por el punto $i = N_1 - 1$, $j = N_2 - 1$ bien sea por las filas, o bien por las columnas de la red al de crecer el índice correspondiente i o j. Como resultado y_{k+1} se determina por las fórmulas

$$y_{k+1}(i, j) = y_k(i, j) - \tau_{k+1} w_k(i, j),$$
$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1, \tag{34}$$
$$y_{k+1}|\gamma = g.$$

Aquí se emplean las siguientes notaciones:

$$\alpha = \frac{\omega_0 h_2^2}{h_1^2 h_2^2 + \omega_0 (h_1^2 + h_2^2)}, \quad \beta = \frac{\omega_0 h_1^2}{h_1^2 h_2^2 + \omega_0 (h_1^2 + h_2^2)},$$
$$\varkappa = \frac{h_1^2 h_2^2}{h_1^2 h_2^2 + \omega_0 (h_1^2 + h_2^2)}. \tag{35}$$

Ya que α, β, $\varkappa > 0$ y $\alpha + \beta + \varkappa = 1$, entonces el cálculo por las fórmulas (32) y (33) es estable. Un cómputo elemental de las operaciones aritméticas para el algoritmo (31)—(35) da $Q_+ = 10 (N_1 - 1)(N_2 - 1)$ operaciones de adición y resta y $Q_* = Q_+$ operaciones de multiplicación, y en total $Q = 20 (N_1 - 1)(N_2 - 1)$.

Tomando en consideración la estimación hallada anteriormente para el número de iteraciones, obtendremos que, para calcular la solución del problema de diferencias (27) mediante el algoritmo (31)—(35) con exactitud ε se debe gastar en el caso $N_1 = N_2 = N$, $l_1 = l_2 = l$

$$Q(\varepsilon) \approx 5{,}6 N^3 \sqrt{N} \ln (2/\varepsilon)$$

operaciones aritméticas.

Examinemos ahora el primer algoritmo, el cual tiene en este caso la forma

$$\varphi_k = (E + \omega_0 R_1)(E + \omega_0 R_2) y_k - \tau_{k+1} (A y_k - f), \tag{36}$$
$$(E + \omega_0 R_1) v = \varphi_k, \quad (E + \omega_0 R_2) y_{k+1} = v.$$

En este algoritmo al pasar a su escritura en diferencias nos será cómodo trabajar con funciones reticulares, definidas sobre $\bar{\omega}$ y que se anulan sobre γ. Estas funciones coinciden con y_k, v y y_{k+1} sobre ω y, como es usual, se designan por $\overset{\bullet}{y}_k$, $\overset{\bullet}{v}$ y $\overset{\bullet}{y}_{k+1}$. Para obtener una aproximación a la solución del problema (27), ella se debe definir así : $y_k(x) = \overset{\circ}{y}_k(x)$ para $x \in \omega$ y $y_k(x) = g(x)$, $x \in \gamma$.

Para realizar en (36) el paso a la escritura en diferencias (puntualmente), es necesario definir el operador de diferencias $\overline{\mathscr{R}}$, el cual corresponde al producto de operadores R_1 y R_2. Notemos que en virtud de la definición los operadores R_1 y R_2 se escriben de la siguiente forma:

$$R_1 y = \begin{cases} \frac{1}{h_1} y_{\bar{x}_1} + \frac{1}{h_2} y_{\bar{x}_2}, & 2 \leqslant i \leqslant N_1 - 1,\ 2 \leqslant j \leqslant N_2 - 1, \\ \frac{1}{h_1^2} y + \frac{1}{h_2} y_{\bar{x}_2}, & i = 1,\ 2 \leqslant j \leqslant N_2 - 1, \\ \frac{1}{h_1} y_{\bar{x}_1} + \frac{1}{h_2^2} y, & 2 \leqslant i \leqslant N_1 - 1,\ j = 1, \\ \left(\frac{1}{h_1^2} + \frac{1}{h_2^2}\right) y, & i = j = 1. \end{cases}$$

$$R_2 y = \begin{cases} -\frac{1}{h_1} y_{x_1} - \frac{1}{h_2} y_{x_2}, & 1 \leqslant i \leqslant N_1 - 2,\ 1 \leqslant j \leqslant N_2 - 2, \\ \frac{1}{h_1^2} y - \frac{1}{h_2} y_{x_2}, & i = N_1 - 1,\ 1 \leqslant j \leqslant N_2 - 2, \\ -\frac{1}{h_1} y_{x_1} + \frac{1}{h_2^2} y, & 1 \leqslant i \leqslant N_1 - 2,\ j = N_2 - 1, \\ \left(\frac{1}{h_1^2} + \frac{1}{h_2^2}\right) y, & i = N_1 - 1,\ j = N_2 - 1. \end{cases}$$

Los razonamientos muestran que si el operador $\overline{\mathscr{R}}$ se define de la forma siguiente:

$$\overline{\mathscr{R}} y = \sum_{\alpha=1}^{2} \frac{1}{h_\alpha^2} y_{\bar{x}_\alpha x_\alpha} + \frac{1}{h_1 h_2} (y_{x_2 \bar{x}_1} + y_{x_1 \bar{x}_2} + qy,\ x \in \omega,$$

donde

$$q(i,\ j) = \begin{cases} 0, & 2 \leqslant i \leqslant N_1 - 1,\ 2 \leqslant j \leqslant N_2 - 1, \\ \frac{1}{h_1^4}, & i = 1,\ 2 \leqslant j \leqslant N_2 - 1, \\ \frac{1}{h_2^4}, & 2 \leqslant i \leqslant N_1 - 1,\ j = 1, \\ \frac{1}{h_1^4} + \frac{1}{h_2^4}, & i = 1,\ j = 1, \end{cases}$$

entonces $R_1 R_2 y = -\overline{\mathscr{R}} \overset{\circ}{y}$, donde $y \in \overset{\circ}{H}$, $\overset{\circ}{y} \in H$ e $y(x) = \overset{\circ}{y}(x)$ para $x \in \omega$.

Utilicemos la definición de los operadores A, R_1 y R_2, y también la expresión obtenida para R_1R_2 y escribamos el algoritmo (36) en la forma

$$\varphi_k(i, j) = [d_{k+1} - q(i, j)\omega_0^2]\overset{\circ}{y}_k(i, j) + \\ + a_{k+1}[\overset{\bullet}{y}_k(i+1, j) + \overset{\circ}{y}_k(i-1, j)] + \\ + b_{k+1}[\overset{\circ}{y}_k(i, j+1) + \overset{\circ}{y}_k(i, j-1)] + \\ + c[\overset{\circ}{y}_k(i-1, j+1) + \overset{\circ}{y}_k(i+1, j-1)] + \tau_{k+1}f(i, j), \quad (37)$$

donde está designado

$$a_{k+1} = \frac{\tau_{k+1}}{h_1^2} - \frac{\omega_0}{h_1^2}\left(1 + \frac{\omega_0}{h_1^2} + \frac{\omega_0}{h_2^2}\right),$$

$$b_{k+1} = \frac{\tau_{k+1}}{h_2^2} - \frac{\omega_0}{h_2^2}\left(1 + \frac{\omega_0}{h_1^2} + \frac{\omega_0}{h_2^2}\right),$$

$$c = \frac{\omega_0^2}{h_1^2 h_2^2}, \quad d_{k+1} = 1 - 2(a_{k+1} + b_{k+1} + c).$$

A continuación

$$v(i, j) = \alpha v(i-1, j) + \beta v(i, j-1) + \varkappa\varphi_k(i, j),$$
$$i = 1, 2, \ldots, N_1 - 1, \; j = 1, 2, \ldots, N_2 - 1, \quad (38)$$
$$v(0, j) = 0, \quad 1 \leqslant j \leqslant N_2 - 1, \quad v(i, 0) = 0, \quad 1 \leqslant i \leqslant N_1 - 1,$$

$$\overset{\circ}{y}_{k+1}(i, j) = \alpha\overset{\circ}{y}_{k+1}(i+1, j) + \beta\overset{\circ}{y}_{k+1}(i, j+1) + \varkappa v(i, j),$$
$$i = N_1 - 1, \; N_1 - 2, \ldots, 1, \; j = N_2 - 1, \; N_2 - 2, \ldots, 1, \quad (39)$$

$$\overset{\bullet}{y}_{k+1}|_\gamma = 0,$$

donde α, β y $\varkappa$ están definidos en (35). El cálculo del número de operaciones aritméticas da $Q_+ = [11N_1N_2] - 10(N_1 + N_2) + 10$ operaciones de adición y resta y $Q_* = Q_+$ operaciones de multiplicación, y en total $Q = 22N_1N_2 - 20(N_1 + N_2) + 20$. Esto es aproximadamente en 1,1 veces mayor, que en el algoritmo (31)—(34). La ventaja del algoritmo (37)—(39) consiste en que aquí no se exige memoria complementaria para conservar la información intermedia $\varphi_k(i, j)$, $v(i, j)$ y la aproximación nuevamente definida $\overset{\bullet}{y}_{k+1}(i, j)$ se ubican sucesivamente en el lugar que ocupaba $y_k(i, j)$.

OBSERVACIÓN 2. En el caso de p dimensiones el operador $\mathscr{R}$ tiene la forma

$$\mathscr{R}y = \sum_{\alpha=1}^{p} \frac{1}{h_\alpha^2} y_{\bar{x}_\alpha x_\alpha} + \sum_{\alpha=1}^{p} \sum_{\beta \neq \alpha}^{1+p} \frac{1}{h_\alpha^2 h_\beta^2} y_{x_\beta \bar{x}_\alpha} + qy,$$

donde

$$q(i_1, i_2, \ldots, i_p) = \sum_{\alpha=1}^{p} \frac{\delta_{l_\alpha, 1}}{h_\alpha^4}, \quad \delta_{i,j} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

OBSERVACION 3. Al método alternativo triangular de bloques le corresponde la siguiente definición de los operadores de diferencias $\mathscr{R}_1$ y $\mathscr{R}_2$:

$$\mathscr{R}_1 y = -\frac{1}{2} y_{\bar{x}_1 x_2} - \frac{1}{h_2} y_{\bar{x}_2}, \quad \mathscr{R}_2 y = -\frac{1}{2} y_{\bar{x}_2 x_1} + \frac{1}{h_2} y_{x_2}.$$

En este caso para invertir el operador B es necesario utilizar el método de factorización tripuntual. Esto conduce al aumento de trabajo computacional en una iteración, lo cual no compensa la pequeña disminución del número de iteraciones (aproximadamente en 1,2 veces).

§ 2. Problemas de contorno de diferencias para ecuaciones elípticas en un rectángulo

1. **Problema de Dirichlet para la ecuación de coeficientes variables.** Examinemos ahora una aplicación del método alternativo triangular a encontrar la solución del problema de Dirichlet de diferencias para la ecuación elíptica sin derivadas mixtas

$$\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha(x) y_{\bar{x}_\alpha})_{x_\alpha} = -\varphi(x), \quad x \in \omega, \tag{1}$$

$$y(x) = g(x), \quad x \in \gamma,$$

en un rectángulo, donde $\bar{\omega} = \omega \cup \gamma$ es una red rectangular uniforme con pasos h_1 y h_2: $\bar{\omega} = \{x_{ij} = (ih_1, jh_2),\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2\}$. Supondremos, que los coeficientes $a_\alpha(x)$ satisfacen las condiciones

$$0 < c_1 \leqslant a_\alpha(x) \leqslant c_2, \quad \alpha = 1, 2. \tag{2}$$

Exigiremos también, que para j fijo $1 \leqslant j \leqslant N_2 - 1$, el número de nodos de la red ω, en los cuales $(a_1)_{x_1} = O(h_1^{-1})$, sea finito y no dependa de h_1. Esto significa que el coeficiente correspondiente en la ecuación diferencial posee un número finito de puntos de discontinuidad por la dirección x_1 para cada x_2 fijo. Una exigencia análoga debe ser cumplida también para $(a_2)_{x_2}$.

El problema de diferencias (1) se reduce a la ecuación operacional

$$Au = f \tag{3}$$

de la manera usual. Aquí H es el espacio de las funciones reticulares, definidas sobre ω, con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x) v(x) h_1 h_2,$$

$$Ay = -\Lambda \overset{\circ}{y},\ y \in H,\ \overset{\circ}{y} \in \overset{\circ}{H} \text{ y } y(x) = \overset{\circ}{y}(x) \text{ para } x \in \omega;$$

$$f(x) = \varphi(x) + \frac{1}{h_1^2} \varphi_1(x) + \frac{1}{h_2^2} \varphi_2(x),$$

donde

$$\varphi_1(x) = \begin{cases} a_1(h_1, x_2)\, g(0, x_2), & x_1 = h_1, \\ 0, & 2h_1 \leqslant x_1 \leqslant l_1 - 2h_1, \\ a_1(l_1, x_2)\, g(l_1, x_2), & x_1 = l_1 - h_1, \end{cases}$$

$$\varphi_2(x) = \begin{cases} a_2(x_1, h_2)\, g(x_1, 0), & x_2 = h_2, \\ 0, & 2h_2 \leqslant x_2 \leqslant l_2 - 2h_2, \\ a_2(x_1, l_2)\, g(x_1, l_2), & x_2 = l_2 - h_2. \end{cases}$$

Utilizando las fórmulas de Green de diferencias, hallaremos, que el operador A es autoconjugado en H y tiene lugar la igualdad

$$(Ay, y) = -(\Lambda \overset{\circ}{y}, \overset{\circ}{y}) = \sum_{\alpha=1}^{2} (a_\alpha \overset{\circ}{y}^2_{\bar{x}_\alpha}, 1)_\alpha, \tag{4}$$

donde

$$(u, v)_\alpha = \sum_{x_\alpha = h_\alpha}^{l_\alpha} \sum_{x_\beta = h_\beta}^{l_\beta - h_\beta} u(x) v(x) h_\alpha h_\beta,$$

$$\beta = 3 - \alpha, \quad \alpha = 1, 2.$$

Para la resolución aproximada de la ecuación examinaremos el método alternativo triangular, construido para la utilización de un regularizador $R \neq A$:

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f, \quad k=0,\ 1,\ \ldots,\ y_0\in H, \qquad (5)$$

$$B=(E+\omega R_1)(E+\omega R_2), \quad R_1=R_2^*,$$

$$R=R_1+R_2.$$

El regularizador R lo eligiremos de la manera siguiente:

$$Ry=-\mathscr{R}\overset{\circ}{y}, \quad \mathscr{R}y=y_{\bar{x}_1x_1}+y_{\bar{x}_2x_2}, \quad \overset{\circ}{y}\in\overset{\circ}{H}, \qquad (6)$$

y definiremos los operadores R_1 y R_2 por las fórmulas

$$R_\alpha(y)=-\mathscr{R}_\alpha\overset{\circ}{y}, \quad \mathscr{R}_1y=-\sum_{\alpha=1}^{2}\frac{1}{h_\alpha}y_{\bar{x}_\alpha}, \quad \mathscr{R}_2y=$$

$$=\sum_{\alpha=1}^{2}\frac{1}{h_\alpha}y_{x_\alpha}. \qquad (7)$$

En el punto 4 del § 1 fue mostrado, que para los operadores R_1 y R_2 aquí definidos tienen lugar las desigualdades

$$\delta E\leqslant R, \quad R_1R_2\leqslant(\Delta/4)R,$$

$$\delta=\sum_{\alpha=1}^{2}\frac{4}{h_\alpha^2}\operatorname{sen}^2\frac{\pi h_\alpha}{2l_\alpha}, \quad \Delta=\sum_{\alpha=1}^{2}\frac{4}{h_\alpha^2}.$$

A continuación, utilizando las fórmulas de Green de diferencias, obtenemos

$$(Ry,\ y)=-(\mathscr{R}\overset{\circ}{y},\ \overset{\circ}{y})=\sum_{\alpha=1}^{2}(\overset{\circ}{y}_{\bar{x}_\alpha}^2,\ 1)_\alpha. \qquad (8)$$

Por consiguiente, de (2), (4) y (8) se infieren las desigualdades $c_1R\leqslant A\leqslant c_2R$, $c_1>0$. Ya que el operador R es autoconjugado y definido positivo en H, entonces para el método examinado tiene lugar el teorema 2 con $\mathscr{D}=E$ en el cual está indicada la elección de los parámetros iterativos ω y $\{\tau_k\}$. De este teorema obtenemos la estimación para el error

$$\|z_n\|_D\leqslant q_n\|z_0\|_D, \quad D=A,\ B\ \text{o}\ AB^{-1}A,$$

donde

$$q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{c_1}{c_2}\frac{2\sqrt{\eta}}{1+\sqrt{\eta}} \quad \eta = \frac{\delta}{\Delta}.$$

Para η pequeño obtendremos una estimación para el número de iteraciones:

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \sqrt{\frac{c_2}{c_1}}\frac{l_n(2/\varepsilon)}{2\sqrt{2}\sqrt[4]{\eta}}, \quad \eta = \frac{\pi^2}{4N^2}.$$

De aquí se deduce que el número de iteraciones es proporcional a $\sqrt{c_2/c_1}$ y es racional utilizar el método (5), (7), cuando esta relación no es demasiado grande.

2. **Método alternativo triangular modificado** *). Continuemos el estudio del método alternativo triangular para el problema de diferencias (1) en el caso, cuando los coeficientes $a_\alpha(x)$ varían fuertemente, es decir, la relación c_2/c_1 es grande.

Examinaremos ahora para la ecuación (3) una variante modificada del método alternativo triangular

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \tag{9}$$

$$B = (\mathscr{D} + \omega R_1)\mathscr{D}^{-1}(\mathscr{D} + \omega R_2), \quad R_1 = R_2^*, \quad R_1 + R_2 = A,$$

donde ponemos $\mathscr{D}y = d(x)y$, $x \in \omega$. Aquí $d(x)$ es cierta función reticular positiva sobre ω sujeta a definición. En este caso $\mathscr{D}$ es un operador autoconjugado y definido positivo en H. La función reticular $d(x)$ juega en (9) el papel de un parámetro iterativo complementario y permite tener en cuenta las singularidades del operador A en cada nodo x de la red ω.

Definamos ahora los operadores R_α de la forma siguiente: $R_\alpha y = -R_\alpha \overset{\cdot}{y}$, $y \in H$, $\overset{\cdot}{y} \in \overset{\cdot}{H}$, donde

$$\mathscr{R}_1 y = -\sum_{\alpha=1}^{2}\left(\frac{a_\alpha}{h_\alpha}y_{\bar{x}_\alpha} + \frac{a_{\alpha x_\alpha}}{2h_\alpha}y\right), \tag{10}$$

$$\mathscr{R}_2 y = \sum_{\alpha=1}^{2}\left(\frac{a_\alpha^{+1}}{h_\alpha}y_{x_\alpha} + \frac{a_{\alpha x_\alpha}}{2h_\alpha}y\right), \quad x \in \omega,$$

$y\ a_1^{\pm 1}(x) = a_1(x_1 \pm h_1, x_2)$, $a_2^{\pm 1}(x) = a_2(x_1, x_2 \pm h_2)$.

*) Véase A. B. Kucherov y E. S. Nikolaiev. (RMS y FM, 16, No. 5, 1976; 17, No. 3, 1977.)

Mostremos que los operadores R_1 y R_2 son conjugados en H. Para esto es suficiente mostrar que tiene lugar la igualdad $(\mathring{\mathscr{R}}_1\mathring{y}, \mathring{v}) = (\mathring{y}, \mathring{\mathscr{R}}_2\mathring{v})$, $\mathring{y} \in \mathring{H}$, $\mathring{v} \in \mathring{H}$. De las fórmulas de Green de diferencias para las funciones que se anulan sobre γ, y de la fórmula de diferenciación de diferencias del producto de funciones reticulares $(yv)_{x_\alpha} = y^{+1}v_{x_\alpha} + y_{x_\alpha} v$ se deduce que

$$(\mathring{\mathscr{R}}_1\mathring{y}, \mathring{v}) = -\sum_{\alpha=1}^{2} \frac{1}{h_\alpha}(a_\alpha \mathring{y}_{\bar{x}_\alpha}, \mathring{v}) - \sum_{\alpha=1}^{2} \frac{1}{2h_\alpha}(a_{\alpha x_\alpha}\mathring{y}, \mathring{v}) =$$

$$= \sum_{\alpha=1}^{2}\left[\frac{1}{h_\alpha}(\mathring{y}, (a_\alpha \mathring{v})_{x_\alpha}) - \frac{1}{2h_\alpha}(a_{\alpha x_\alpha}\mathring{y}, \mathring{v})\right] =$$

$$= \sum_{\alpha=1}^{2}\left[\frac{1}{h_\alpha}(\mathring{y}, a_\alpha^{+1}\mathring{v}_{x_\alpha}) + \frac{1}{2h_\alpha}(\mathring{y}, a_{\alpha x_\alpha}\mathring{v}) = (\mathring{y}, \mathring{\mathscr{R}}_2\mathring{v}).\right.$$

La afirmación está demostrada.

Puesto que $R_1 + R_2 = A$, entonces en virtud del teorema 1 la información a priori para el método alternativo triangular (9) posee la forma de las constantes δ y Δ de las desigualdades

$$\delta\mathscr{D} \leqslant A, \quad R_1\mathscr{D}^{-1}R_2 \leqslant \frac{\Delta}{4}A, \quad \delta > 0. \tag{11}$$

Como la relación $\eta = \frac{\delta}{\Delta}$ determina el número de iteraciones, entonces la función reticular $d(x)$ debe ser elegida de la condición de maximalidad de esta relación.

Ocupémonos ahora de la elección de la función $d(x)$ y de las estimaciones de δ y Δ. Demostremos primeramente una desigualdad.

LEMA 2. *Sean $p_\alpha(x)$, $q_\alpha(x)$, $u_\alpha(x)$ y $v_\alpha(x)$, $\alpha = 1, 2$ las funciones reticulares definidas sobre ω. Entonces para cualquier $x \in \omega$ tiene lugar la desigualdad*

$$\left[\sum_{\alpha=1}^{2}(p_\alpha u_\alpha + q_\alpha v_\alpha)\right]^2 \leqslant$$

$$\leqslant (1+\varepsilon)(|p_1| + \varkappa_1|q_1|)\left(|p_1|u_1^2 + \frac{|q_1|}{\varkappa_1}v_1^2\right) +$$

$$+ \frac{1+\varepsilon}{\varepsilon}(|p_2| + \varkappa_2|q_2|)\left(|p_2|u_2^2 + \frac{|q_2|}{\varkappa_2}v_2^2\right), \tag{12}$$

donde $\varepsilon(x)$, $\varkappa_1(x)$ *y* $\varkappa_2(x)$ *son las funciones reticulares arbitrarias positivas sobre* ω.

En efecto, utilizando ε, la desigualdad $2ab \leqslant \varepsilon a^2 + b^2/\varepsilon$, $\varepsilon > 0$, obtendremos

$$\Big[\sum_{\alpha=1}^{2}(p_\alpha u_\alpha + q_\alpha v_\alpha)\Big]^2 = (p_1u_1 + q_1v_1)^2 + 2(p_1u_1 + q_1v_1)(p_2u_2 + q_2v_2) + (p_2u_2 + q_2v_2)^2 \leqslant$$
$$\leqslant (1+\varepsilon)(p_1u_1 + q_1v_1)^2 + \frac{1+\varepsilon}{\varepsilon}(p_2u_2 + q_2v_2)^2. \qquad (13)$$

Utilizando de nuevo la desigualdad indicada, hallaremos

$$(p_\alpha u_\alpha + q_\alpha v_\alpha)^2 = p_\alpha^2 u_\alpha^2 + 2p_\alpha q_\alpha u_\alpha v_\alpha + q_\alpha^2 v_\alpha^2 \leqslant$$
$$\leqslant p_\alpha^2 u_\alpha^2 + |p_\alpha|\,|q_\alpha|\left(\varkappa_\alpha u_\alpha^2 + \frac{1}{\varkappa_\alpha} v_\alpha^2\right) + q_\alpha^2 v_\alpha^2 =$$
$$= (|p_\alpha| + \varkappa_\alpha |q_\alpha|)\left(|p_\alpha| u_\alpha^2 + \frac{|q_\alpha|}{\varkappa_\alpha} v_\alpha^2\right), \quad \varkappa_\alpha > 0, \quad \alpha = 1, 2.$$

Sustituyendo la desigualdad obtenida en (13), tendremos (12). El lema está demostrado.

Aprovechemos la desigualdad (12) y también la definición de los operadores R_1 y R_2 y hallaremos, que

$$(R_1\mathscr{D}^{-1}R_2 y,\ y) = (\mathscr{D}^{-1}R_2\overset{\circ}{y},\ R_2\overset{\circ}{y}) =$$
$$= \Big(\frac{1}{d}\sum_{\alpha=1}^{2}\Big(\frac{a_\alpha^{+1}}{h_\alpha}\overset{\circ}{y}_{x_\alpha} + \frac{a_\alpha x_\alpha}{2h_\alpha}\overset{\circ}{y}\Big)^2,\ 1\Big) \leqslant$$
$$\leqslant \Big(\frac{(1+\varepsilon)}{dh_1^2}(a_1^{+1} + 0{,}5h_1\varkappa_1 |a_{1x_1}|)\Big(a_1^{+1}\overset{\circ}{y}_{x_1}^2 + \frac{0{,}5h_1|a_{1x_1}|}{\varkappa_1 h_1^2}\overset{\circ}{y}^2\Big),\ 1\Big) +$$
$$+ \Big(\frac{(1+\varepsilon)}{d\varepsilon h_2^2}(a_2^{+1} + 0{,}5h_2\varkappa_2 |a_{2x_2}|)\Big(a_2^{+1}\overset{\circ}{y}_{x_2}^2 + \frac{0{,}5h_2|a_{2x_2}|}{\varkappa_2 h_2^2}\overset{\circ}{y}^2\Big),\ 1\Big).$$

Senalemos que en (12) en lugar de p_α, q_α, u_α y v_α pusimos $p_2 = \frac{a_\alpha^{+1}}{h_\alpha}$, $q_\alpha = 0{,}5a_{\alpha x_\alpha}$, $u_\alpha = \overset{\circ}{y}_{x_\alpha}$, $v_\alpha = \frac{1}{h_\alpha}\overset{\circ}{y}$, $\alpha = 1, 2$. Exigiremos que en la desigualdad obtenida $\varkappa_1$ sea la función solamente de x_2, y $\varkappa_2$ únicamente de x_1, es decir, pondremos

$$\varkappa_\alpha = \varkappa_\alpha(x_\beta), \quad \beta = 3 - \alpha, \quad \alpha = 1, 2. \qquad (14)$$

Pongamos

$$\varepsilon = \varepsilon(x) = \frac{a_2^{+1} + 0{,}5h_2\varkappa_2 |a_{2x_2}|}{a_1^{+1} + 0{,}5h_1\varkappa_1 |a_{1x_1}|} \cdot \frac{h_1^2\theta_2(x_1)}{h_2^2\theta_1(x_2)} \tag{15}$$

y definamos $d(x)$ de la forma siguiente:

$$d(x) = \sum_{\alpha=1}^{2} (a_\alpha^{+1} + 0{,}5h_\alpha\varkappa_\alpha |a_{\alpha x_\alpha}|) \frac{\theta_\alpha}{h_\alpha^2}, \tag{16}$$

donde $\theta_\alpha = \theta_\alpha(x_\beta)$, $\beta = 3 - \alpha$, $\alpha = 1, 2$, son las funciones reticulares positivas sobre ω sujetas a definición.

Sustituyendo (15) y (16) en la desigualdad obtenida anteriormente, tendremos

$$(R_1\mathscr{D}^{-1}R_2y, y) \leqslant \sum_{\alpha=1}^{2} \frac{a_\alpha^{+1}}{\theta_\alpha} \overset{\circ}{y}{}^2_{x_\alpha}, 1) + \sum_{\alpha=1}^{2} \left(\frac{|a_{\alpha x_\alpha}|}{2h_\alpha\theta_\alpha\varkappa_\alpha} \overset{*}{y}{}^2, 1 \right).$$

Ya que θ_α no depende de x_α, entonces, utilizando el producto escalar $(,)_\alpha$ introducido anteriormente, obtendremos que

$$\left(\frac{a_\alpha^{+1}}{\theta_\alpha} \overset{*}{y}{}^2_{x_\alpha}, 1 \right) \leqslant \left(\frac{a_\alpha}{\theta_\alpha} \overset{\circ}{y}{}^2_{x_\alpha}, 1 \right)_\alpha, \quad \alpha = 1, 2.$$

Por consiguiente,

$$(R_1\mathscr{D}^{-1}R_2y, y) \leqslant \sum_{\alpha=1}^{2} \left(\frac{a_\alpha}{\theta_\alpha} \overset{\circ}{y}{}^-_{x_\alpha}, 1 \right)_\alpha + \sum_{\alpha=1}^{2} \left(\frac{|a_{\alpha x_\alpha}|}{2h_\alpha\theta_\alpha\varkappa_\alpha} \overset{\circ}{y}{}^2, 1 \right). \tag{17}$$

Elijamos ahora θ_α y $\varkappa_\alpha$. Designemos

$$\omega_1 = \{x_1 = ih_1, \quad 1 \leqslant i \leqslant N_1 - 1, \ h_1N_1 = l_1\},$$
$$\omega_1^+ = \{x_1 = ih_1, \quad 1 \leqslant i \leqslant N_1, \quad h_1N_1 = l_1\}$$

y definamos

$$(u, v)_{\omega_1} = \sum_{x_1 \in \omega_1} u(x) v(x) h_1,$$
$$(u, v)_{\omega_1^+} = \sum_{x_1 \in \omega_1^+} u(x) v(x) h_1.$$

De modo análogo se introducen ω_2 y ω_2^+, y también $(u, v)_{\omega_2}$ y $(u, v)_{\omega_2^+}$. Entonces es fácil ver, que tienen lugar las relaciones

$$(u, v) = ((u, v)_{\omega_1}, 1)_{\omega_2} = ((u, v)_{\omega_2}, 1)_{\omega_1},$$
$$(u, v)_\alpha = ((u, v)_{\omega_\alpha^+}, 1)_{\omega_\beta}, \quad \beta = 3 - \alpha, \quad \alpha = 1, 2. \tag{18}$$

Sea ahora $b_\alpha(x_\beta) = \max\limits_{x_\alpha \in \omega_\alpha} v^\alpha(x)$, $\alpha = 1, 2$, $x_\beta \in \omega_\beta$, donde $v^\alpha(x)$ para x_β fijo es la solución del siguiente problema de contorno tripuntual:

$$(a_\alpha v^\alpha_{\overline{x}_\alpha})_{x_\alpha} = -\frac{a_\alpha^{+1}}{h_\alpha^2}, \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \tag{19}$$
$$v^\alpha(x) = 0, \quad x_\alpha = 0, l_\alpha, \quad x_\beta \in \omega_\beta.$$

Entonces en virtud del lema 13 del punto 4 del § 2 del cap. V obtenemos

$$\left(\frac{a_\alpha^{+1}}{h_\alpha^2} \overset{\circ}{y}{}^2, 1\right)_{\omega_\alpha} \leqslant b_\alpha(x_\beta)(a_\alpha \overset{\circ}{y}{}^2_{\overline{x}_\alpha}, 1)_{w_\alpha^+}, \quad \alpha = 1, 2.$$

Multipliquemos esta desigualdad por $\theta_\alpha(x_\beta)$ y adicionemos escalarmente por ω_β. Entonces en virtud de (18) tendremos

$$\left(\frac{\theta_\alpha a_\alpha^{+1}}{h_\alpha^2} \overset{\circ}{y}{}^2, 1\right) \leqslant (b_\alpha a_\alpha \theta_\alpha \overset{\circ}{y}{}^2_{\overline{x}_\alpha}, 1)_\alpha, \quad \alpha = 1, 2. \tag{20}$$

Sea $c_\alpha(x_\beta) = \max\limits_{x_\alpha \in \omega_\alpha} \omega^\alpha(x)$, $\alpha = 1, 2$, $x_\beta \in \omega_\beta$, donde $\omega^\alpha(x)$ para x_β fijo es la solución del siguiente problema de contorno tripuntual:

$$(a_\alpha \omega^\alpha_{\overline{x}_\alpha})_{x_\alpha} = -\frac{|a_{\alpha x_\alpha}|}{2h_\alpha}, \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \tag{21}$$
$$w^\alpha(x) = 0, \quad x_\alpha = 0, l_\alpha, \quad x_\beta \in \omega_\beta.$$

Análogamente a como fueron obtenidas las desigualdades (20), en virtud de (14) tendremos las desigualdades siguientes:

$$\left(\frac{\varkappa_\alpha \theta_\alpha |a_{\alpha x_\alpha}|}{2h_\alpha} \overset{\circ}{y}{}^2, 1\right) \leqslant (\varkappa_\alpha \theta_\alpha c_\alpha a_\alpha \overset{\circ}{y}{}^2_{\overline{x}_\alpha}, 1)_\alpha, \quad \alpha = 1, 2, \tag{22}$$

$$\left(\frac{|a_{\alpha x_\alpha}|}{2h_\alpha \theta_\alpha \varkappa_\alpha} \overset{\circ}{y}{}^2, 1\right) \leqslant \left(\frac{c_\alpha}{\varkappa_\alpha \theta_\alpha} a_\alpha \overset{\circ}{y}{}^2_{\overline{x}_\alpha}, 1\right)_\alpha, \quad \alpha = 1, 2. \tag{23}$$

Adicionemos ahora las desigualdades (20) y (22) y sumémoslas por α. Entonces en virtud de (16) obtendremos

$$(\overset{\bullet}{dy}^2,\ 1) = (\mathscr{D}y,\ y) \leqslant ((\varkappa_\alpha c_\alpha + b_\alpha)\,\theta_\alpha a_\alpha \overset{\circ}{y}{}^2_{\bar{x}_\alpha},\ 1)_\alpha.$$

Eligiendo θ_α según la fórmula

$$\theta_\alpha(x_\beta) = \frac{1}{b_\alpha(x_\beta) + c_\alpha(x_\beta)\,\varkappa_\alpha(x_\beta)},\ \beta = 3-\alpha,\ \alpha = 1,\ 2, \qquad (24)$$

y teniendo en cuenta (4), de aquí hallaremos que $(\mathscr{D}y,\ y) \leqslant (Ay,\ y)$. Por lo tanto, en (11) se puede poner $\delta = 1$.

Estimemos ahora Δ. Para eso sustituyamos (23) en (17) y tengamos en cuenta la elección de θ_α por la fórmula (24). Como resultado obtendremos la estimación siguiente:

$$(R_1\mathscr{D}^{-1}R_2 y,\ y) \leqslant \sum_{\alpha=1}^{2} ((1 + c_\alpha/\varkappa_\alpha)(b_\alpha + c_\alpha\varkappa_\alpha)\,a_\alpha \overset{\circ}{y}{}^2_{\bar{x}_\alpha},\ 1)_\alpha.$$

Elijamos ahora el $\varkappa_\alpha$ óptimo de la condición de mínimo según $\varkappa_\alpha$ de la expresión $(1 + c_\alpha/\varkappa_\alpha)(b_\alpha + c_\alpha\varkappa_\alpha)$. Obtendremos $\varkappa_\alpha(x_\beta) = \sqrt{b_\alpha(x_\beta)}$, $\beta = 3 - \alpha$, $\alpha = 1,\ 2$ y con esto

$$(R_1\mathscr{D}^{-1}R_2 y,\ y) \leqslant \sum_{\alpha=1}^{2} ((c_\alpha + \sqrt{b_\alpha})^2\,a_\alpha \overset{\circ}{y}{}^2_{\bar{x}_\alpha},\ 1)_\alpha.$$

Comparando esta estimación con (4), hallaremos que en las desigualdades (11) se puede poner

$$\Delta = 4 \max_{\alpha=1,\,2} \left(\max_{x_\beta \in \omega_\beta} (c_\alpha(x_\beta) + \sqrt{b_\alpha(x_\beta)})^2 \right),\ \beta = 3-\alpha. \qquad (25)$$

Sustituyendo las expresiones obtenidas para $\varkappa_\alpha$ y θ_α en (16), obtenemos para la función $d(x)$ una representación del tipo

$$d(x) = \sum_{\alpha=1}^{2} \left(\frac{a_\alpha^{+1}}{h_\alpha^2 \sqrt{b_\alpha}} + \frac{|a_{\alpha x_\alpha}|}{2h_\alpha} \right) \frac{1}{c_\alpha + \sqrt{b_\alpha}},\quad x \in \omega. \qquad (26)$$

De esta forma, hemos hallado la función $d(x)$ y las constantes δ y Δ. Ahora queda por aplicar el teorema 1. Señalemos que como $\delta = 1$, entonces $\omega_0 = 2/\sqrt{\Delta}$, y para el número de iteraciones es válida la estimación

$$n \geqslant n_0(\varepsilon),\quad n_0(\varepsilon) = \frac{\sqrt[4]{\Delta}\ln(2/\varepsilon)}{2\sqrt{2}}.$$

A continuación, en virtud de las condiciones (2) de (19) y (21) obtenemos que $b_\alpha = O(1/h_\alpha^2)$ y $c_\alpha = O(1/h_\alpha)$, si el número de puntos en los cuales $a_{\alpha x_\alpha} = O(h_\alpha^{-1})$, es finito. De aquí se deduce que $n_0(\varepsilon) = O(\sqrt{N} \ln(2/\varepsilon))$.

Detengámonos ahora en la realización de la variante construida del método alternativo triangular (9). Al principio para x_β fija, $h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta$, se resuelven por el método de factorización los problemas de contorno tripuntuales (19) y (21) y se encuentran los valores de $b_\alpha(x_\beta)$ y $c_\alpha(x_\beta)$, $\alpha = 1, 2$. Estas cuatro funciones reticulares unidimensionales se memorizan y se utilizan en el proceso de las iteraciones para calcular $d(x)$ por la fórmula (26). La sencillez de la fórmula (26) permite no conservar la función reticular bidimensional $d(x)$, y calcularla de nuevo a medida que sea necesario.

Después, por la fórmula (25) se encuentra Δ y se pone $\delta = 1$. Los valores de los parámetros iterativos ω y τ_k para el esquema (9) se determinan de acuerdo con el teorema 1.

Para encontrar la aproximación y_{k+1} por y_k dada se utiliza el primero de los algoritmos descritos para el método alternativo triangular en el punto 1 del § 1:

$$(\mathscr{D} + \omega_0 R_1) v = \varphi_k, \quad (\mathscr{D} + \omega_0 R_2) y_{k+1} = \mathscr{D} v,$$
$$\varphi_k = (\mathscr{D} + \omega_0 R_1) \mathscr{D}^{-1} (\mathscr{D} + \omega_0 R_2) y_k - \tau_{k+1} \times \quad (27)$$
$$\times (A y_k - f).$$

Sin detenernos en detalles citaremos la forma de diferencias del algoritmo (27):

$$v(i, j) = \alpha_1(i, j) v(i-1, j) + \beta_1(i, j) v(i, j-1) +$$
$$+ \varkappa(i, j) \varphi_k(i, j),$$
$$i = 1, 2, \ldots, N_1 - 1, \; j = 1, 2, \ldots, N_2 - 1, \quad (28)$$
$$v(0, j) = 0, \quad 1 \leqslant j \leqslant N_2 - 1, \quad v(i, 0) = 0,$$
$$1 \leqslant i \leqslant N_1 - 1.$$

$$\overset{\bullet}{y}_{k+1}(i, j) = \alpha_2(i, j) \overset{\bullet}{y}_{k+1}(i+1, j) +$$
$$+ \beta_2(i, j) \overset{\bullet}{y}_{k+1}(i, j+1) + \varkappa(i, j) d(i, j) v(i, j),$$
$$i = N_1 - 1, \ldots, 1, \quad j = N_2 - 1, \ldots, 1, \quad (29)$$

donde

$$\alpha_1 = \frac{\omega_0 a_1 \varkappa}{h_1^2}, \quad \beta_1 = \frac{\omega_0 a_2 \varkappa}{h_2^2}, \quad \alpha_2 = \frac{\omega_0 a_1^{+1} \varkappa}{h_1^2}, \quad \beta_2 = \frac{\omega_0 a_2^{+1} \varkappa}{h_2^2},$$

$$\frac{1}{\varkappa} = d + \omega_0 \left[\frac{a_1^{+1} + a_1}{2h_1^2} + \frac{a_2^{+1} + a_2}{2h_2^2} \right].$$

El segundo miembro $\varphi_h(i, j)$ se calcula mediante las fórmulas

$$\varphi_h(i, j) = [P(i-1, j) + Q(i, j-1) + S(i, j)] \times$$
$$\times \overset{\circ}{y}_h(i, j) + R_1(i, j) \overset{\circ}{y}_h(i+1, j) + R_1(i-1, j) \times$$
$$\times \overset{\circ}{y}_h(i-1, j) + R_2(i, j) \overset{\circ}{y}_h(i, j+1) + R_2(i, j-1) \times$$
$$\times \overset{\circ}{y}_h(i, j-1) + G(i-1, j) \overset{\circ}{y}_h(i-1, j+1) +$$
$$+ G(i, j-1) \overset{\circ}{y}_h(i+1, j-1) + \tau_{h+1} f(i, j), \quad (30)$$
$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1,$$

donde

$$G = \frac{\omega_0^2 a_1^{+1} a_2^{+1}}{h_1^2 h_2^2 d}, \quad R_\alpha = \left(\tau_{h+1} - \frac{\omega_0}{d\varkappa} \right) \frac{a_\alpha^{+1}}{h_\alpha^2}, \quad \alpha = 1, 2,$$

$$S = \frac{1}{\omega_0 \varkappa} \left[\frac{\omega_0}{d\varkappa} - 2\tau_{h+1}(1 - \varkappa d) \right],$$

$$P = \frac{\omega_0^2 (a_1^{+1})^2}{h_1^4 d}, \quad Q = \frac{\omega_0^2 (a_2^{+1})^2}{h_2^4 d},$$

y al mismo tiempo $P(0, j) = 0$, $1 \leqslant j \leqslant N_2 - 1$, $Q(i, 0) = 0$, $1 \leqslant i \leqslant N_1 - 1$. Notemos que en virtud de (25) y (26) son ciertas las estimaciones

$$c_\alpha + \sqrt{b_\alpha} \leqslant \frac{\sqrt{\Delta}}{2} = \frac{1}{\omega_0}, \quad d \geqslant \omega_0 \sum_{\alpha=1}^{2} \frac{|a_{\alpha x_\alpha}|}{2h_\alpha}.$$

De aquí obtenemos

$$\frac{1}{\varkappa} = d + \omega_0 \sum_{\alpha=1}^{2} \frac{a_\alpha^{+1} + a_\alpha}{2h_\alpha^2} \geqslant \omega_0 \sum_{\alpha=1}^{2} \left(\frac{|a_{\alpha x_\alpha}|}{2h_\alpha} + \frac{a_\alpha^{+1} + a_\alpha}{2h_\alpha^2} \right) =$$
$$= \omega_0 \left(\text{máx} \left(\frac{a_1^{+1}}{h_1^2}, \frac{a_1}{h_1^2} \right) + \text{máx} \left(\frac{a_2^{+1}}{h_2^2}, \frac{a_2}{h_2^2} \right) \right)$$

ó

$$\text{máx}(\alpha_1, \alpha_2) + \text{máx}(\beta_1, \beta_2) \leqslant 1.$$

De aquí se deduce, que $\alpha_1 + \beta_1 \leqslant 1$ y $\alpha_2 + \beta_2 \leqslant 1$. Por eso el cálculo según las fórmulas (28) y (30) es estable.

3. **Comparación de las variantes del método.** Más arriba para resolver el problema de diferencias (1) fueron construidas dos variantes del método alternativo triangular. La variante (5), (7) está construida a base del regularizador R, mientras la variante (9), (10) utiliza un operador $\mathscr{D}$, el cual se elige de una manera especial. Estas variantes se caracterizan por una misma dependencia asintótica entre el número de iteraciones y el número de nodos de la red. Sin embargo la estimación del número de iteraciones para la primera variante depende de las características extremales de los coeficientes $a_\alpha(x)$, $\alpha = 1, 2$, de la ecuación en diferencias (1), a la vez que para la segunda variante se determina por sus características integrales.

Comparemos estas variantes del método en el siguiente ejemplo de modelo tradicional. Sea dada la ecuación en diferencias (1) sobre una red cuadrada con $N_1 = N_2 = N$ introducida en el cuadrado unidad ($l_1 = l_2 = 1$), en la cual

$$a_1(x) = 1 + c\,[(x_1 - 0{,}5)^2 + (x_2 - 0{,}5)^2],$$
$$a_2(x) = 1 + c\,[0{,}5 - (x_1 - 0{,}5)^2 - (x_2 - 0{,}5)^2],$$
$$x \in \bar{\omega}.$$

Entonces en las desigualdades (2) tenemos $c_1 = 1$, $c_2 = 1 + 0{,}5c$. Cambiando el parámetro c, obtendremos los coeficientes $a_\alpha(x)$ con diferentes propiedades extremales.

Tabla 10

c_2/c_1	$N=32$		$N=64$		$N=128$	
	(5), (7)	(9), (10)	(5), (7)	(9), (10)	(5), (7)	(9), (10)
2	23	18	32	26	45	36
8	46	21	64	30	90	43
32	92	23	128	34	180	49
128	184	24	256	36	360	53
512	367	24	512	36	720	54

En la tabla 10 se cita el número de iteraciones para las variantes, indicadas en función del número N de nodos por una dirección y de la relación c_2/c_1 para $\varepsilon = 10^{-4}$. Se ve que para los valores grandes de c_2/c_1 el método alternativo

triangular modificado exige de un menor número de iteraciones y al mismo tiempo el número de iteraciones depende debilmente de esta relación.

4. **Tercer problema de contorno.** Examinemos el método alternativo triangular para resolver el tercer problema de contorno para una ecuación elíptica en el rectángulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$:

$$\sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha}\left(k_\alpha(x)\frac{\partial u}{\partial x_\alpha}\right) = -\varphi(x), \quad x \in G,$$

$$k_\alpha \frac{\partial u}{\partial x_\alpha} = \varkappa_{-\alpha}(x)\,u - g_{-\alpha}(x), \quad x_\alpha = 0, \tag{31}$$

$$-k_\alpha \frac{\partial u}{\partial x_\alpha} = \varkappa_{+\alpha}(x)\,u - g_{+\alpha}(x), \quad x_\alpha = l_\alpha.$$

Sobre la red rectangular $\overline{\omega} = \{x_{ij} = (ih_1, jh_2) \in \overline{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2\}$ al problema (31) le corresponde el problema de diferencias

$$\Lambda y = -f(x), \quad x \in \overline{\omega},$$
$$\Lambda = \Lambda_1 + \Lambda_2, \quad f(x) = \varphi(x) + \frac{2}{h_1}\varphi_1(x) + \frac{2}{h_2}\varphi_2(x), \tag{32}$$

donde

$$\Lambda_\alpha y = \begin{cases} \dfrac{2}{h_\alpha}\left(a_\alpha^{+1} y_{\bar{x}_\alpha} - \varkappa_{-\alpha} y\right), & x_\alpha = 0, \\ \left(a_\alpha y_{\bar{x}_\alpha}\right)_{x_\alpha}, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ \dfrac{2}{h_\alpha}\left(-a_\alpha y_{\bar{x}_\alpha} - \varkappa_{+\alpha} y\right), & x_\alpha = l_\alpha, \end{cases}$$

$$\varphi_\alpha(x) = \begin{cases} q_{-\alpha}(x_\alpha), & x_\alpha = 0, \\ 0, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ g_{+\alpha}(x_\alpha), & x_\alpha = l_\alpha. \end{cases}$$

Supondremos que los coeficientes $a_\alpha(x)$ satisfacen las condiciones (2) y poseen un número finito de puntos, en los cuales $a_{\alpha x_\alpha} = O(h_\alpha^{-1})$. También consideramos que $\varkappa_{-\alpha}(x_\beta)$ y $\varkappa_{+\alpha}(x_\beta)$ para cada x_β fija no se anulan simultáneamente ($\varkappa_{-\alpha} \geqslant 0$, $\varkappa_{+\alpha} \geqslant 0$, $\varkappa_{-\alpha} + \varkappa_{+\alpha} > 0$).

Es cómodo reducir previamente el problema de diferencias (32) al problema de Dirichlet en la región ampliada $\overline{\omega}^* = \{x_{ij} = (ih_1, jh_2),\ -1 \leqslant i \leqslant N_1 + 1,\ -1 \leqslant j \leqslant N_2 +$

$+1\}$, para la cual la red $\bar{\omega}$ es interior. Designemos mediante γ^* la frontera de la red $\bar{\omega}^*$ y definamos la función reticular $y(x)$ igual a cero sobre γ^*. Si designamos

$$\bar{a}_\alpha(x) = \begin{cases} \rho(x_\beta)\, h_\alpha \varkappa_{-\alpha}(x_\beta), & x_\alpha = 0, \\ \rho(x_\beta)\, a_\alpha(x), & h_\alpha \leqslant x_\alpha \leqslant l_\alpha, \\ \rho(x_\beta)\, h_\alpha \varkappa_{+\alpha}(x_\beta), & x_\alpha = l_\alpha + h_\alpha, \quad 0 \leqslant x_\beta \leqslant l_\beta \end{cases}$$

$$\bar{f}(x) = \rho(x_1)\,\rho(x_2)\, f(x), \quad x \in \bar{\omega},$$

$$\rho(x_\beta) = \begin{cases} 0{,}5, & x_\beta = 0, \quad l_\beta, \\ 1, & h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta, \end{cases}$$

$$\beta = 3 - \alpha, \quad \alpha = 1, 2,$$

entonces el problema (32) puede ser escrito en la forma

$$\bar{\Lambda} y = \sum_{\alpha=1}^{2} (\bar{a}_\alpha y_{\bar{x}_\alpha})_{x_\alpha} = -\bar{f}(x), \quad x \in \bar{\omega}, \tag{33}$$

$$y(x) = 0, \quad x \in \gamma^*.$$

Recordemos que para el problema de diferencias del tipo (33) en el punto 2 fue construido el método alternativo triangular modificado (9)—(10). Por consiguiente, en las fórmulas del punto 2 es preciso sólo cambiar $a_\alpha(x)$ por $\bar{a}_\alpha(x)$, y nosotros obtendremos un método para resolver el tercer problema de contorno para una ecuación elíptica en un rectángulo.

Para el caso a examinar el problema de contorno tripuntual (19) se escribe en la forma

$$(\bar{a}_\alpha v^\alpha_{\bar{x}_\alpha})_{x_\alpha} = -\frac{\bar{a}^{+1}_\alpha}{h^2_\alpha}, \quad 0 \leqslant x_\alpha \leqslant l_\alpha,$$

$$v^\alpha(x) = 0, \quad x_\alpha = -h_\alpha, \quad l_\alpha + h_\alpha. \tag{34}$$

Utilizando las notaciones introducidas más arriba para $\bar{a}_\alpha$, obtendremos que a (34) se le puede dar otra forma

$$(a_\alpha v^\alpha_{\bar{x}_\alpha})_{x_\alpha} = -\frac{a^{+1}_\alpha}{h^2_\alpha}, \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,$$

$$a^{+1}_\alpha v^\alpha_{\bar{x}_\alpha} - \varkappa_{-\alpha} v^\alpha = -\frac{a^{+1}_\alpha}{h_\alpha}, \quad x_\alpha = 0, \tag{35}$$

$$-a_\alpha v^\alpha_{\bar{x}_\alpha} - \varkappa_{+\alpha} v^\alpha = -\varkappa_{+\alpha}, \quad x_\alpha = l_\alpha.$$

En virtud de las suposiciones hechas respecto a $a_\alpha(x)$, $\varkappa_{-\alpha}$ y $\varkappa_{+\alpha}$, el problema de diferencias es soluble. Con esto

$$b_\alpha(x_\beta) = \max_{0 \leqslant x_\alpha \leqslant l_\alpha} v^\alpha(x) = O\left(\frac{1}{h_\beta^2}\right), \quad 0 \leqslant x_\beta \leqslant l_\beta.$$

Análogamente el problema (21), el cual en este caso tiene la forma

$$(\bar{a}_\alpha w^\alpha_{\bar{x}_\alpha})_{x_\alpha} = -\frac{|\bar{a}_{\alpha x_\alpha}|}{2h_\alpha}, \quad 0 \leqslant x_\alpha \leqslant l_\alpha,$$
$$w^\alpha(x) = 0, \quad x_\alpha = -h_\alpha, \quad l_\alpha + h_\alpha,$$

en virtud de las notaciones se reduce al tercer problema de contorno

$$(a_\alpha w^\alpha_{\bar{x}_\alpha})_{x_\alpha} = -\frac{|a_{\alpha x_\alpha}|}{2h_\alpha}, \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,$$
$$a^{+1}_\alpha w^\alpha_{\bar{x}_\alpha} - \varkappa_{-\alpha} w^\alpha = -\frac{|a^{+1}_\alpha - h_\alpha \varkappa_{-\alpha}|}{2h_\alpha}, \quad x_\alpha = 0, \tag{36}$$
$$-a_\alpha w^\alpha_{\bar{x}_\alpha} - \varkappa_{+\alpha} w^\alpha = -\frac{|a_\alpha - h_\alpha \varkappa_{+\alpha}|}{2h_\alpha}, \quad x_\alpha = l_\alpha.$$

De aquí obtenemos, que

$$c_\alpha(x_\beta) = \max_{0 \leqslant x_\alpha \leqslant l_\alpha} w^\alpha(x) = O\left(\frac{1}{h_\alpha}\right),$$

y por lo tanto,

$$\Delta = 4 \max_{\alpha=1,2} \left(\max_{0 \leqslant x_\beta \leqslant l_\beta} (c_\alpha(x_\beta) + \sqrt{b_\alpha(x_\beta)})^2 \right) = O\left(\frac{1}{|h|^2}\right).$$

Por eso para el método alternativo triangular modificado, aplicado a encontrar la solución del tercer problema de contorno (32), el número de iteraciones depende del número de nodos igual que en el caso del primer problema de contorno.

Es evidente, que el método de reducción al problema de Dirichlet descrito más arriba se puede utilzar también en el caso, cuando sobre cada lado del rectángulo está prefijada una de las condiciones de contorno de primero, de segundo o de tercer género.

5. Problema de Dirichlet en diferencias para una ecuación de derivadas mixtas. Supongamos que en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \alpha = 1, 2\}$ con frontera Γ se exige hallar

la solución del problema de Dirichlet para la ecuación de tipo elíptico con derivadas mixtas

$$Lu = \sum_{\alpha,\ \beta=1}^{2} \frac{\partial}{\partial x_\alpha}\left(k_{\alpha\beta}(x)\frac{\partial u}{\partial x_\beta}\right) = -\varphi(x), \quad x \in G, \quad (37)$$

$$u(x) = g(x), \quad x \in \Gamma, \quad k_{\alpha\beta}(x) = k_{\beta\alpha}(x).$$

Exigiremos el cumplimiento de las condiciones

$$k_{\alpha\alpha}(x) \geqslant c > 0, \quad k_{12}^2(x) \leqslant \rho^2 k_{11}(x)\, k_{22}(x), \quad x \in \overline{G}, \quad (38)$$

$$0 \leqslant \rho < 1.$$

Señalemos que las condiciones (38) aseguran la elipticidad uniforme de la ecuación (37). En efecto, examinemos para $x \in G$ fija el problema en valores propios para el haz de matrices

$$\begin{Vmatrix} k_{11} & k_{12} \\ k_{12} & k_{22} \end{Vmatrix} - \lambda \begin{Vmatrix} k_{11} & 0 \\ 0 & k_{22} \end{Vmatrix} = 0.$$

Para λ tenemos la ecuación cuadrática $(1-\lambda)^2 k_{11}k_{22} - k_{12}^2 = 0$. De aquí hallaremos

$$|1-\lambda| = \frac{|k_{12}|}{\sqrt{k_{11}k_{22}}} \leqslant \rho, \quad 1-\rho \leqslant \lambda \leqslant 1+\rho.$$

Por consiguiente, tiene lugar la desigualdad

$$c_1 \sum_{\alpha=1}^{2} |k_{\alpha\alpha}(x)\, \xi_\alpha^2 \leqslant \sum_{\alpha,\ \beta=1}^{2} k_{\alpha\beta}(x)\, \xi_\alpha \xi_\beta \leqslant c_2 \sum_{\alpha=1}^{2} k_{\alpha\alpha}(x)\, \xi_\alpha^2, \quad (39)$$

$$c_1 = 1 - \rho, \quad c_2 = 1 + \rho,$$

donde $\xi = (\xi_1, \xi_2)$ es un vector arbitrario. De aquí en virtud de la condición $k_{\alpha\alpha}(x) \geqslant c > 0$ se sigue la elipticidad uniforme del operador L.

Sobre la red rectangular $\overline{\omega} = \{x_{ij} = (ih_1, jh_2) \in \overline{G}, 0 \leqslant i \leqslant N_1, 0 \leqslant j \leqslant N_2, h_\alpha N_\alpha = l_\alpha, \alpha = 1, 2\}$ le pondremos en correspondencia al problema (37), (38) el problema de Dirichlet de diferencias

$$\Lambda y = \frac{1}{2} \sum_{\alpha,\ \beta=1}^{2} [(k_{\alpha\beta} y_{\bar{x}_\beta})_{x_\alpha} + (k_{\alpha\beta} y_{x_\beta})_{\bar{x}_\alpha}] = -\varphi(x), \quad x \in \omega, \quad (40)$$

$$y(x) = g(x), \quad x \in \gamma.$$

En el espacio H de las funciones reticulares, definidas sobre ω, con el producto escalar

$$(u,\ v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2$$

definiremos el operador A de la siguiente forma: $Ay = -\Delta \overset{\circ}{y}$, $y \in H$ e $y(x) = \overset{\circ}{y}(x)$ para $x \in \omega$, $\overset{\circ}{y}(x) = 0$ para $x \in \gamma$, y también el operador R: $Ry = -\mathscr{R}\overset{\circ}{y}$, donde

$$\mathscr{R}y = \sum_{\alpha=1}^{2} (a_\alpha y_{\bar{x}_\alpha})_{x_\alpha}, \quad x \in \omega,\ a_\alpha(x) = \frac{k_{\alpha\alpha} + k^{-1}_{\alpha\alpha}}{2}.$$

Entonces el problema (40) se puede escribir en forma de la ecuación (3), donde $f(x)$ se diferencia de $\varphi(x)$ únicamente en los nodos fronterizos. Puesto que $k_{\alpha\beta}(x) = k_{\beta\alpha}(x)$, entonces los operadores A y R son autoconjugados. Mostremos que tienen lugar las desigualdades

$$c_1 R \leqslant A \leqslant c_2 R, \quad c_1 = 1 - \rho, \quad c_2 = 1 + \rho, \qquad (41)$$

donde ρ está prefijada en (38). En efecto, de las fórmulas de Green de diferencias obtenemos

$$(Ay,\ y) = -(\Lambda \overset{\circ}{y},\ \overset{\circ}{y}) = \sum_{\alpha,\ \beta=1}^{2} \frac{1}{2}\Big[(k_{\alpha\beta}\overset{\circ}{y}_{x_\beta},\ \overset{\circ}{y}_{\bar{x}_\alpha})_\alpha + {}_\alpha(k_{\alpha\beta}\overset{\circ}{y}_{x_\beta},\ \overset{\circ}{y}_{x_\alpha})\Big],$$

donde el producto escalar $(u,\ v)$ está definido en el punto 1 del § 2, y

$${}_\alpha(u,\ v) = \sum_{x_\alpha=0}^{l_\alpha - h_\alpha} \sum_{x_\beta=h_\beta}^{l_\beta - h_\beta} u(x)\, v(x)\, h_1 h_2, \quad \beta = 3-\alpha, \quad \alpha = 1,\ 2.$$

Notemos que si una de las funciones $u(x)$ o $v(x)$ se anula para $x_\beta = 0$ (o para $x_\beta = l_\beta$), entonces

$${}_\alpha(u,\ v) = [u,\ v) = \sum_{x_1=0}^{l_1-h_1} \sum_{x_2=0}^{l_2-h_2} u(x)\, v(x)\, h_1 h_2,$$

$$(u,\ v)_\alpha = (u,\ v] = \sum_{x_1=h_1}^{l_1} \sum_{x_2=h_2}^{l_2} u(x)\, v(x)\, h_1 h_2, \quad \alpha = 1,\ 2.$$

Esto inmediatamente da

$$(Ay,\ y)=\sum_{\alpha,\ \beta=1}^{2}\frac{1}{2}\{(k_{\alpha\beta}\overset{\circ}{y}_{\bar{x}_\beta},\ \overset{\circ}{y}_{\bar{x}_\alpha}]+[k_{\alpha\beta}\overset{\circ}{y}_{x_\beta},\ \overset{\circ}{y}_{x_\alpha})\}. \qquad (42)$$

A continuación, como $\mathcal{R}y$ se puede escribir en la forma

$$\mathcal{R}y=\frac{1}{2}\sum_{\alpha=1}^{2}\{(k_{\alpha\alpha}y_{\bar{x}_\alpha})_{x_\alpha}+(k_{\alpha\alpha}y_{x_\alpha})_{\bar{x}_\alpha}\},$$

entonces

$$(Ry,\ y)=-(\mathcal{R}\overset{\circ}{y},\ \overset{\circ}{y})=\sum_{\alpha=1}^{2}\frac{1}{2}\{(k_{\alpha\alpha}\overset{\bullet}{y}_{\bar{x}_\alpha},\ \overset{\circ}{y}_{\bar{x}_\alpha})_\alpha+$$

$$+{}_\alpha(k_{\alpha\alpha}\overset{\circ}{y}_{x_\alpha},\ \overset{\bullet}{y}_{x_\alpha})\}=$$

$$=\sum_{\alpha=1}^{2}\frac{1}{2}\{(k_{\alpha\alpha}\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1]+[k_{\alpha\alpha}\overset{\circ}{y}^2_{x_\alpha},\ 1)\}. \qquad (43)$$

De (42), (43) y de las desigualdades (39) obtenemos

$$c_1\Big(\sum_{\alpha=1}^{2}k_{\alpha\alpha}\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1\Big]\leqslant\Big(\sum_{\alpha,\ \beta=1}^{2}k_{\alpha\beta}\overset{\bullet}{y}_{\bar{x}_\beta}\overset{\circ}{y}_{\bar{x}_\alpha},\ 1\Big]\leqslant$$

$$\leqslant c_2\Big(\sum_{\alpha=1}^{2}k_{\alpha\alpha}\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1\Big]$$

y análogamente

$$c_1\Big[\sum_{\alpha=1}^{2}k_{\alpha\alpha}\overset{\circ}{y}^2_{\bar{x}_\alpha},\ 1\Big)\leqslant\Big[\sum_{\alpha,\ \beta=1}^{2}k_{\alpha\beta}\overset{\circ}{y}_{x_\beta}\overset{\bullet}{y}_{x_\alpha},\ 1\Big)\leqslant c_2\Big[\sum_{\alpha=1}^{2}k_{\alpha\alpha}\overset{\circ}{y}^2_{x_\alpha},\ 1\Big),$$

y por lo tanto, las estimaciones (41) están demostradas.

De esta manera, el operador R definido más arriba, se puede utilizar en calidad de regularizador en el método alternativo triangular

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f,\quad k=0,\ 1,\ \ldots,$$

$$B=(\mathcal{D}+\omega R_1)\,\mathcal{D}^{-1}(\mathcal{D}+\omega R_2),\quad R_1=R_1^*,\quad R_1+R_2=R,$$

donde los operadores R_1, R_2 y $\mathcal{D}$ están definidos en el punto 2 del § 2. Allí fueron halladas las constantes δ y Δ para las

desigualdades $\delta\mathscr{D} \leqslant R$, $R_1\mathscr{D}^{-1}R_2 \leqslant \frac{\Delta}{4}R$, $\delta > 0$. La aplicación del teorema 2 concluye la construcción del método alternativo triangular para el problema de diferencias (40).

§ 3. Método alternativo triangular para ecuaciones elípticas en una región arbitraria

1. Planteamiento del problema de diferencias. Construyamos un método alternativo triangular modificado para resolver el problema de Dirichlet en una región acotada arbitraria $\bar{G}$ con frontera Γ en el caso de una ecuación elíptica de coeficientes variables

$$\sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha}\left(k_\alpha(x)\frac{\partial u}{\partial x_\alpha}\right) = -\varphi(x), \quad x \in G, \tag{1}$$

$$u(x) = g(x), \quad x \in \Gamma, \quad k_\alpha(x) \geqslant c_1 > 0, \quad \alpha = 1, 2.$$

Supongamos, que la frontera Γ es suficientemente suave. Además, para simplificar la exposición consideraremos que la intersección de la región con una recta que pasa por todo punto $x \in G$ paralelamente al eje de coordenadas Ox_α, $\alpha = 1, 2$ consiste de un intervalo.

En la región $\bar{G}$ construiremos de modo siguiente una red no uniforme $\bar{\omega}$. Tracemos la familia de rectas $x_\alpha = x_\alpha(i_\alpha)$, $i_\alpha = 0, \pm 1, \pm 2, \ldots$, $\alpha = 1, 2$. Entonces los puntos $x_i = (x_1(i_1), x_2(i_2))$, $i = (i_1, i_2)$ forman una red fundamental sobre el plano. A un punto x_i de la red, que pertenezca a G, lo llamaremos nodo interior de la red $\bar{\omega}$. Al conjunto de todos los nodos interiores, como es usual, lo designaremos mediante ω.

La intersección de cualquier recta que pase por el punto $x_i \in \omega$ paralelamente al eje Ox_α con la región G es el intervalo $\Delta_\alpha(x_1)$. Los extremos de este intervalo los llamaremos nodos de frontera por la dirección x_α. El conjunto de todos los nodos frontera por x_α lo designaremos mediante γ_α. La frontera de la red $\bar{\omega}$ es $\gamma = \gamma_1 \cup \gamma_2$, de modo que $\bar{\omega} = \omega \cup \gamma$. La red $\bar{\omega}$ está construida.

Introduzcamos una serie de notaciones. Designemos mediante $\omega_\alpha(x_\beta)$, $\beta = 3 - \alpha$, $\alpha = 1, 2$ el conjunto de los

nodos de la red ω que se encuentran sobre el intervalo Δ_α; mediante $\omega_\alpha^+(x_\beta)$ el conjunto que consta de $\omega_\alpha(x_\beta)$ y del extremo derecho del intervalo Δ_α; $\overline{\omega}_\alpha(x_\beta)$ que consiste de $\omega_\alpha(x_\beta)$ y los extremos del intervalo Δ_α.

Designemos por $x^{(+1\alpha)}$ y $x^{(-1\alpha)}$ los nodos vecinos de $x \in \omega_\alpha(x_\beta)$ a la derecha y a la izquierda y que pertenecen a $\overline{\omega}_\alpha(x_\beta)$. Notemos que si, por ejemplo, $x^{(+1\alpha)} \in \gamma_\alpha$, entonces este nodo puede no coincidir con un nodo de la red fundamental.

Definamos $h_\alpha^+(x) = x^{(+1\alpha)} - x$, $h_\alpha^-(x) = x - x^{(-1\alpha)}$, $x \in \omega_\alpha$, $x^{(\pm 1\alpha)} \in \overline{\omega}_\alpha$. En todos los nodos interiores de la red ω definiremos también los pasos medios $\hbar_\alpha(x_\alpha) = 0{,}5\,(x_\alpha \times \times (i_\alpha + 1) - x_\alpha(i_\alpha - 1)$ como la distancia entre las correspondientes rectas de la red fundamental.

Al problema (1) le pondremos en correspondencia sobre la red $\overline{\omega}$ el problema de diferencias

$$\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha y_{\bar{x}_\alpha})_{\hat{x}_\alpha} = -\varphi(x), \quad x \in \omega, \tag{2}$$

$$y(x) = g(x), \quad x \in \gamma.$$

Aquí son utilizadas las notaciones siguientes:

$$(a_\alpha y_{\bar{x}_\alpha})_{\hat{x}_\alpha} = \frac{1}{\hbar_\alpha}(a_\alpha^{+1} y_{x_\alpha} - a_\alpha y_{\bar{x}_\alpha}), \quad a_\alpha^{+1} = a_\alpha(x^{(+1\alpha)}),$$

$$y_{x_\alpha} = \frac{1}{h_\alpha^+}(y(x^{(+1\alpha)}) - y(x)), \quad y_{\bar{x}_\alpha} = \frac{1}{h_\alpha^-}(y(x) - y(x^{(-1\alpha)})).$$

Los coeficientes $a_\alpha(x)$ y $\varphi(x)$ son elegidos de manera tal, que el esquema (2) sobre una red uniforme tenga el segundo orden local de aproximación.

Introduzcamos ahora el espacio H de funciones reticulares definidas sobre ω, con el producto escalar $(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, \hbar_1(x_1)\, \hbar_2(x_2)$. Definiremos el operador A de la forma usual: $Ay = -\Lambda \overset{\circ}{y}$, $y \in H$ e $y(x) = \overset{\circ}{y}(x)$ para $x \in \omega$, $\overset{\circ}{y}(x) = 0$ para $x \in \gamma$. Entonces el problema de diferencias (2) se escribe en forma de la ecuación

$$Au = f, \tag{3}$$

donde $f(x)$ se diferencia de $\varphi(x)$ solamente en los nodos fronterizos.

2. Construcción del método alternativo triangular. Para la ecuación (3) examinaremos el método alternativo triangular modificado

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f, \quad k=0, 1, \ldots, \tag{4}$$

$$B=(\mathscr{D}+\omega R_1)\,\mathscr{D}^{-1}(\mathscr{D}+\omega R_2),$$

$$R_1=R_1^*, \quad R_1+R_2=A.$$

Definamos los operadores R_1, R_2 y $\mathscr{D}$. Como en el caso de un rectángulo, eligiremos al operador $\mathscr{D}$ más sencillo:

$$\mathscr{D}y=d(x)\,y, \quad d(x)>0 \quad \text{para} \quad x\in\omega, \tag{5}$$

y pondremos $R_\alpha y=-\mathscr{R}_\alpha \mathring{y}$, $y\in H$ e $y(x)=\mathring{y}(x)$, $x\in\omega$, donde

$$\mathscr{R}_1 y=-\sum_{\alpha=1}^{2}\left[\frac{a_\alpha}{\hbar_\alpha}y_{\bar{x}_\alpha}+\frac{1}{2\hbar_\alpha}\left(\frac{a_\alpha^{+1}}{h_\alpha^+}-\frac{a_\alpha}{h_\alpha^-}\right)y\right],$$

$$\mathscr{R}_2 y=\sum_{\alpha=1}^{2}\left[\frac{a_\alpha^{+1}}{\hbar_\alpha}y_{x_\alpha}+\frac{1}{2\hbar_\alpha}\left(\frac{a_\alpha^{+1}}{h_\alpha^+}-\frac{a_\alpha}{h_\alpha^-}\right)y\right]. \tag{6}$$

Ya que $\mathscr{R}_1+\mathscr{R}_2=\Lambda$, entonces obtenemos que $R_1+R_2=A$.

Mostremos que los operadores R_1 y R_2 son conjugados. Primeramente introduciremos las notaciones que utilizaremos en lo sucesivo. Definamos las siguientes sumas:

$$(u, v)_{\omega_\alpha(x_\beta)}=\sum_{x_\alpha\in\omega_\alpha(x_\beta)}u(x)\,v(x)\,\hbar_\alpha(x_\alpha),$$

$$(u, v)_{\omega_\alpha^+(x_\beta)}=\sum_{x_\alpha\in\omega_\alpha^+(x_\beta)}u(x)\,v(x)\,h_\alpha^-(x),$$

$$(u, v)_\alpha=((u, v)_{\omega_\alpha^+(x_\beta)}, 1)_{\omega_\beta(x_\alpha)}=$$

$$=\sum_{x_\beta\in\omega_\beta}\sum_{x_\alpha\in\omega_\alpha^+}u(x)\,v(x)\,h_\alpha^-(x)\,\hbar_\beta(x_\beta),$$

$$\beta=3-\alpha, \quad \alpha=1, 2.$$

Utilizando las notaciones introducidas el producto escalar en H se puede escribir en la forma siguiente:

$$(u, v)=((u, v)_{\omega_1(x_2)}, 1)_{\omega_2(x_1)}=((u, v)_{\omega_2(x_1)}, 1)_{\omega_1(x_2)}. \tag{7}$$

Primeramente demostremos una afirmación auxiliar. Supongamos que y_i y v_i son las funciones reticulares, definidas para $0 \leqslant i \leqslant N$, a la vez $y_0 = y_N = 0$ y $v_0 = v_N = 0$. Sea u_i una función reticular, definida para $1 \leqslant i \leqslant N$. Entonces tiene lugar la igualdad

$$\sum_{i=1}^{N-1} (u_{i+1} - u_i)\, y_i v_i = -\sum_{i=1}^{N-1} (v_{i+1} - v_i)\, u_{i+1} y_i - \sum_{i=1}^{N-1} (y_i - y_{i-1})\, u_i v_i. \quad (8)$$

En efecto, tenemos

$$\sum_{i=1}^{N-1} [(u_{i+1} - u_i)\, y_i v_i + (v_{i+1} - v_i)\, u_{i+1} y_i + (y_i - y_{i-1})\, u_i v_i] = \sum_{i=1}^{N-1} (u_{i+1} v_{i+1} y_i - u_i v_i y_{i-1}) = u_N v_N y_{N-1} - u_1 v_1 y_0 = 0.$$

La afirmación (8) está demostrada. Utilizando (8) es fácil mostrar que para las funciones $\overset{\circ}{y}(x)$ y $\overset{\circ}{v}(x)$ definidas sobre $\overline{\omega}$ y que se anulan sobre γ, tiene lugar la igualdad

$$(u_{x_\alpha} \overset{\circ}{y}, \overset{\circ}{v})_{\omega_\alpha} = -\left(\overset{\circ}{y}, \frac{h_\alpha^+}{\hbar_\alpha} u^{+1\alpha} \overset{\circ}{v}_{x_\alpha}\right)_{\omega_\alpha} - \left(\frac{h_\alpha^-}{\hbar_\alpha} u \overset{\circ}{y}_{\overline{x}_\alpha}, \overset{\circ}{v}\right)_{\omega_\alpha}. \quad (9)$$

Aquí son utilizadas las notaciones

$$u^{+1\alpha} = u(x^{(+1\alpha)}), \quad u_{\hat{x}_\alpha} = \frac{u^{+1\alpha} - u}{\hbar_\alpha(x_\alpha)}.$$

Sustituyendo la expresión $u(x) = a_\alpha(x)/h_\alpha^-(x)$ en (9) y teniendo en cuenta la igualdad $h_\alpha^-(x^{(+1\alpha)}) = h_\alpha^+(x)$, obtenemos

$$\left(\frac{1}{\hbar_\alpha}\left(\frac{a_\alpha^{+1}}{h_\alpha^+} - \frac{a_\alpha}{h_\alpha^-}\right)\overset{\circ}{y}, \overset{\circ}{v}\right)_{\omega_\alpha} = -\left(\overset{\circ}{y}, \frac{a_\alpha^{+1}}{\hbar_\alpha} \overset{\circ}{v}_{x_\alpha}\right)_{\omega_\alpha} - \left(\frac{a_\alpha}{\hbar_\alpha} \overset{\circ}{y}_{\overline{x}_\alpha}, \overset{\circ}{v}\right)_{\omega_\alpha}.$$

Multiplicando esta relación por $\hbar_\beta(x_\beta)$, sumando según ω_β y tomando en consideración (7), hallaremos

$$-\left(\frac{a_\alpha}{\hbar_\alpha} \overset{\circ}{y}_{\overline{x}_\alpha}, \overset{\circ}{v}\right) = \left(\frac{a_\alpha^{+1}}{\hbar_\alpha} \overset{\circ}{v}_{x_\alpha}, \overset{\circ}{y}\right) + \left(\frac{1}{\hbar_\alpha}\left(\frac{a_\alpha^{+1}}{h_\alpha^+} - \frac{a_\alpha}{h_\alpha^-}\right)\overset{\circ}{y}, \overset{\circ}{v}\right). \quad (10)$$

Demostremos ahora, que los operadores R_1 y R_2 son conjugados. De (6) y (10), obtenemos

$$(R_1 y, v) = -(\mathscr{R}_1 \overset{\circ}{y}, \overset{\circ}{v}) = \sum_{\alpha=1}^{2} \left[\left(\frac{a_\alpha}{\hbar_\alpha} \overset{\circ}{y}_{\bar{x}_\alpha}, \overset{\circ}{v} \right) + \right.$$

$$\left. + \left(\frac{1}{2\hbar_\alpha} \left(\frac{a_\alpha^{+1}}{h_\alpha^+} - \frac{a_\alpha}{h_\alpha^-} \right) \overset{\circ}{y}, \overset{\circ}{v} \right) \right] = -\sum_{\alpha=1}^{2} \left[\left(\overset{\circ}{y}, \frac{a_\alpha^{+1}}{\hbar_\alpha} v_{x_\alpha} \right) + \right.$$

$$\left. + \left(\frac{1}{2\hbar_\alpha} \left(\frac{a_\alpha^{+1}}{h_\alpha^+} - \frac{a_\alpha}{h_\alpha^-} \right) \overset{\circ}{y}, \overset{\circ}{v} \right) \right] = -(\overset{\circ}{y}, \mathscr{R}_2 \overset{\circ}{v}) = (y, R_2(v)).$$

La afirmación está demostrada. De aquí, a propósito, se deduce también la autoconjugación del operador A.

Nos queda por construir la función $d(x)$, la cual define el operador $\mathscr{D}$, hallar las constantes δ y Δ en las desigualdades

$$\delta \mathscr{D} \leqslant A, \quad R_1 \mathscr{D}^{-1} R_2 \leqslant \frac{\Delta}{4} A, \quad \delta > 0, \tag{11}$$

y aprovecharnos del teorema 1. Todo esto nosotros lo haremos igual que en el punto 2 del § 2, donde fue examinado el problema de Dirichlet para una ecuación elíptica en un rectángulo sobre una red uniforme.

Al principio señalemos que en virtud de las fórmulas de Green de diferencias tiene lugar la igualdad

$$(Ay, y) = \sum_{\alpha=1}^{2} (a_\alpha \overset{\circ}{y}_{\bar{x}_\alpha}, 1)_\alpha, \quad \overset{\circ}{y}(x) = 0, \quad x \in \gamma.$$

Luego de (5) y (6) hallaremos

$$(R_1 \mathscr{D}^{-1} R_2 y, y) = (\mathscr{D}^{-1} \mathscr{R}_2 \overset{\circ}{y}, \mathscr{R}_2 \overset{\circ}{y}) =$$

$$= \left(\frac{1}{d} \sum_{\alpha=1}^{2} \left[\frac{a_\alpha^{+1}}{\hbar_\alpha} \overset{\circ}{y}_{x_\alpha} + \frac{1}{2\hbar_\alpha} \left(\frac{a_\alpha^{+1}}{h_\alpha^+} - \frac{a_\alpha}{h_\alpha^-} \right) y \right]^2, 1 \right).$$

Utilicemos ahora el lema 2, poniendo

$$p_\alpha = \frac{a_\alpha^{+1}}{\hbar_\alpha}, \quad q_\alpha = \frac{h_\alpha^{+1}}{2\hbar_\alpha} \left(\frac{a_\alpha^{+1}}{h_\alpha^+} - \frac{a_\alpha}{h_\alpha^-} \right),$$

$$u_\alpha = \overset{\circ}{y}_{x_\alpha}, \quad v_\alpha = \frac{1}{h_\alpha^+} \overset{\circ}{y}, \quad \alpha = 1, 2.$$

Como resultado obtendremos la desigualdad

$$(R_1\mathscr{D}^{-1}R_2y,\ y)\leqslant\Big(\frac{1+\varepsilon}{d\hbar_1^2}\Big[a_1^{+1}+\frac{\varkappa_1h_1^+}{2}\Big|\frac{a_1^{+1}}{h_1^+}-$$
$$-\frac{a_1}{h_1^-}\Big|\Big]\Big[a_1^{+1}\overset{\circ}{y}{}^2_{x_1}+\frac{1}{2\varkappa_1h_1^+}\Big|\frac{a_1^{+1}}{h_1^+}-\frac{a_1}{h_1^-}\Big|\overset{\circ}{y}{}^2\Big],\ 1\Big)+$$
$$+\Big(\frac{1+\varepsilon}{d\varepsilon\hbar_2^2}\Big[a_2^{+1}+\frac{\varkappa_2h_2^+}{2}\Big|\frac{a_2^{+1}}{h_2^+}-\frac{a_2}{h_2^-}\Big|\Big]\Big[a_2^{+1}\overset{\circ}{y}{}^2_{x_2}+$$
$$+\frac{1}{2\varkappa_2h_2^+}\Big|\frac{a_2^{+1}}{h_2^+}-\frac{a_2}{h_2^-}\Big|\overset{\circ}{y}{}^2\Big],\ 1\Big).$$

Pongamos aquí

$$\varepsilon=\varepsilon\ (x)=\frac{a_2^{+1}+0{,}5\varkappa_2h_2^+\Big|\frac{a_2^{+1}}{h_2^+}-\frac{a_2}{h_2^-}\Big|}{a_1^{+1}+0{,}5\varkappa_1h_1^+\Big|\frac{a_1^{+1}}{h_1^+}-\frac{a_1}{h_1^-}\Big|}\ \frac{\hbar_1h_1^+}{\hbar_2h_2^+}\ \frac{\theta_2}{\theta_1}$$

y definamos $d\,(x)$ de modo siguiente:

$$d\,(x)=\sum_{\alpha=1}^{2}\Big(a_\alpha^{+1}+\frac{\varkappa_\alpha h_\alpha^+}{2}\Big|\frac{a_\alpha^{+1}}{h_\alpha^+}-\frac{a_\alpha}{h_\alpha^-}\Big|\Big)\frac{\theta_\alpha}{\hbar_\alpha h_\alpha^+},\quad x\in\omega.$$

Aquí se supone que $\varkappa_\alpha=\varkappa_\alpha x_\beta>0$, $\theta_\alpha=\theta_\alpha\,(x_\beta)>0$, $\beta=3-\alpha$, $\alpha=1,\ 2$. Como resultado obtendremos la desigualdad

$$(R_1\mathscr{D}^{-1}R_2y,\ y)\leqslant\sum_{\alpha=1}^{2}\Big(\frac{a_\alpha^{+1}h_\alpha^+}{\theta_\alpha\hbar_\alpha}\overset{\circ}{y}{}^2_{x_\alpha},\ 1\Big)+$$
$$+\sum_{\alpha=1}^{2}\Big(\frac{1}{2\hbar_\alpha\theta_\alpha\varkappa_\alpha}\Big|\frac{a_\alpha^{+1}}{h_\alpha^+}-\frac{a_\alpha}{h_\alpha^-}\Big|\overset{\circ}{y}{}^2,\ 1\Big).$$

Puesto que θ_α no depende de x_α, entonces

$$\Big(\frac{a_\alpha^{+1}}{\theta_\alpha}\frac{h_\alpha^+}{\hbar_\alpha}\overset{\circ}{y}{}^2_{x_\alpha},\ 1\Big)\omega_\alpha=\frac{1}{\theta_\alpha}\sum_{x_\alpha\in\omega_\alpha}a_\alpha^{+1}h_\alpha^+\overset{\circ}{y}{}^2_{x_\alpha}\leqslant$$
$$\leqslant\frac{1}{\theta_\alpha}\sum_{x_\alpha\in\omega_\alpha^+}a_\alpha h_\alpha^-\overset{\circ}{y}{}^2_{x_\alpha}.$$

Por lo tanto,

$$\Big(\frac{a_\alpha^{+1}}{\theta_\alpha}\frac{h_\alpha^+}{\hbar_\alpha}\overset{\circ}{y}{}^2_{x_\alpha},\ 1\Big)\leqslant\Big(\frac{a_\alpha}{\theta_\alpha}\overset{\circ}{y}{}^2_{x_\alpha},\ 1\Big)_\alpha,$$

y por eso finalmente tenemos

$$(R_1\mathscr{D}^{-1}R_2y,\ y)\leqslant\sum_{\alpha=1}^{2}\left(\frac{a_\alpha}{\theta_\alpha}\,\overset{\circ}{y}{}^2_{\overline{x}_\alpha},\ 1\right)_\alpha+$$

$$+\sum_{\alpha=1}^{2}\left(\frac{1}{2\hbar_\alpha\theta_\alpha\varkappa_\alpha}\left|\frac{a_\alpha^{+1}}{h_\alpha^+}-\frac{a_\alpha}{h_\alpha^-}\right|\overset{\circ}{y}{}^2,\ 1\right).$$

Los razonamientos ulteriores son completamente análogos a las transformaciones y estimaciones, obtenidas en el punto 2 del § 2. Citemos un resumen: en las desigualdades (11)

$$\delta=1,\ \Delta=4\max_{\alpha=1,\,2}\left(\max_{x_\beta\in\omega_\beta}\left(c_\alpha(x_\beta)+\sqrt{b_\alpha(x_\beta)}\right)^2\right),\quad\beta=3-\alpha,$$

donde

$$b_\alpha(x_\beta)=\max_{x_\alpha\in\omega_\alpha}v^\alpha(x),\quad c_\alpha(x_\beta)=\max_{x_\alpha\in\omega_\alpha}w^\alpha(x),\quad x_\beta\in\omega_\beta;$$

la función $v^\alpha(x)$ es la solución del problema de contorno tripuntual

$$(a_\alpha v^\alpha_{\overline{x}_\alpha})_{\hat{x}_\alpha}=-\frac{a_\alpha^{+1}}{\hbar_\alpha h_\alpha^+},\quad x_\alpha\in\omega_\alpha(x_\beta),\qquad(12)$$
$$v^\alpha(x)=0,\quad x_\alpha\in\gamma_\alpha,$$

y la función $w^\alpha(x)$ es la solución del problema

$$(a_\alpha w^\alpha_{\overline{x}_\alpha})_{\hat{x}_\alpha}=-\frac{1}{2\hbar_\alpha}\left|\frac{a_\alpha^{+1}}{h_\alpha^+}-\frac{a_\alpha}{h_\alpha^-}\right|,\quad x_\alpha\in\omega_\alpha(x_\beta),\qquad(13)$$
$$w^\alpha(x)=0,\quad x_\alpha\in\gamma_\alpha.$$

Con esto la función $d(x)$ se calcula por la fórmula

$$d(x)=\sum_{\alpha=1}^{2}\left(\frac{a_\alpha^{+1}}{\hbar_\alpha h_\alpha^+\sqrt{b_\alpha}}+\frac{1}{2\hbar_\alpha}\left|\frac{a_\alpha^{+1}}{h_\alpha^+}-\right.\right.$$

$$\left.\left.-\frac{a_\alpha}{h_\alpha^-}\right|\right)\frac{1}{c_\alpha+\sqrt{b_\alpha}},\quad x\in\omega.$$

Los parámetros iterativos ω y $\{\tau_k\}$ se calculan mediante las fórmulas del teorema 1. Para encontrar y_{k+1} se puede aprovechar el algoritmo

$$v(x) = \alpha_1(x)\, v^{(-1,1)} + \beta_1(x)\, v^{(-1,2)} + \varkappa(x)\, \varphi_k(x), \quad x \in \omega.$$

$$v(x) = 0, \quad x \in \gamma,$$

$$\overset{\circ}{y}_{k+1}(x) = \alpha_2(x)\, \overset{\circ}{y}{}^{(+1,1)}_{k+1} + \beta_2(x)\, \overset{\circ}{y}{}^{(+1,2)}_{k+1} + \varkappa(x)\, d(x)\, v(x), \quad x \in \omega,$$

$$y_{k+1}(x) = 0, \quad x \in \gamma,$$

donde

$$\alpha_1 = \frac{\omega_0 a_1 \varkappa}{\hbar_1 h_1^-}, \quad \beta_1 = \frac{\omega_0 a_2 \varkappa}{\hbar_2 h_2^-}, \quad \alpha_2 = \frac{\omega_0 a_1^{+1} \varkappa}{\hbar_1 h_1^+}, \quad \beta_2 = \frac{\omega_0 a_1^{+1} \varkappa}{\hbar^2 h_2^+},$$

$$\frac{1}{\varkappa} = d + \frac{1}{2} \sum_{\alpha=1}^{2} \frac{\omega_0}{\hbar_\alpha} \left(\frac{a_\alpha^{+1}}{h_\alpha^+} + \frac{a_\alpha}{h_\alpha^-} \right),$$

$$\varphi_k(x) = By_k - \tau_{k+1}(Ay_k - f).$$

Se debe señalar que en casos semejantes, cuando el cálculo del valor By_k exige grandes gastos de trabajo computacional, y no hay limitaciones para el volumen de información intermedia a memorizar, es racional utilizar el segundo algoritmo, descrito en el punto 1 del § 1.

3. Problema de Dirichlet para la ecuación de Poisson en una región arbitraria. Examinemos en calidad de ejemplo de aplicación del método construido el problema de Dirichlet para la ecuación de Poisson

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -\varphi(x), \quad x \in G, \quad u(x) = g(x), \quad x \in \Gamma.$$

Supondremos que la red es cuadrada, es decir, $x_\alpha = x_\alpha(i_\alpha)$,

$$x_\alpha(i_\alpha + 1) = x_\alpha(i_\alpha) + h, \; i_\alpha = 0, \pm 1, \pm 2, \ldots, \text{ y } \omega = \{x_i = (i_1 h, \; i_2 h) \in G, \; i_\alpha = 0, \pm 1, \pm 2, \ldots\}.$$

Con esto $\hbar_\alpha \equiv h$, y los pasos $h_\alpha^\pm(x)$ son diferentes de h solamente en los nodos fronterizos de la red $\overline{\omega}$.

Aprovechemos el esquema de diferencias (2), en el cual pondremos $a_\alpha(x) \equiv 1$ y $\hbar_\alpha \equiv h$. Para aplicar el método alternativo triangular construido en el punto 2 del § 3, es necesario hallar la solución de los problemas de contorno

tripuntuales (12) y (13), los cuales en este caso tienen la forma

$$\Lambda_\alpha v^\alpha = v^\alpha_{\overline{x}_\alpha \hat{x}_\alpha} = -\frac{1}{h h^+_\alpha}, \quad x_\alpha \in \omega_\alpha(x_\beta),$$

$$v^\alpha(x) = 0, \quad x_\alpha \in \gamma_\alpha, \qquad (14)$$

$$\Lambda_\alpha w^\alpha = w^\alpha_{\overline{x}_\alpha \hat{x}_\alpha} = -\frac{1}{2h}\left|\frac{1}{h^+_\alpha} - \frac{1}{h^-_\alpha}\right|, \qquad (15)$$

$$x_\alpha \in \omega_\alpha(x_\beta), \quad w^{\alpha'}(x) = 0, \quad x_\alpha \in \gamma_\alpha.$$

Examinemos el intervalo Δ_α, que contiene $\omega_\alpha(x_\beta)$, y designemos mediante $l_\alpha(x_\beta)$ y $L_\alpha(x_\beta)$ sus extremos izquierdo y derecho. Entonces $h^-_\alpha(x) \leqslant h$, si x es el nodo de la red ω_α más cercano a l_α, y $h^+_\alpha(x) \leqslant h$, si x es el nodo de la red ω_α más cercano a L_α. Todos los pasos $h^\pm_\alpha$ restantes son iguales al paso fundamental h del retículo.

A su vez el operador Λ_α sobre la red ω_α se escribe con detalle en la forma siguiente:

$$\Lambda_\alpha y = \begin{cases} \frac{1}{h}\left(\frac{y^{+1\alpha} - y}{h} - \frac{y - y^{-1\alpha}}{h^-_\alpha}\right), & x_\alpha = l_\alpha + h^-_\alpha, \\ \frac{1}{h^2}(y^{+1\alpha} - 2y + y^{-1\alpha}), & l_\alpha + h^-_\alpha + h \leqslant \\ & \leqslant x_\alpha \leqslant L_\alpha - h^+_\alpha - h, \\ \frac{1}{h}\left(\frac{y^{+1\alpha} - y}{h^+_\alpha} - \frac{y - y^{-1\alpha}}{h}\right), & x_\alpha = L_\alpha - h^+_\alpha. \end{cases}$$

Las soluciones de las ecuaciones (14), (15) se pueden hallar en forma explícita. Para esto sustituyamos las condiciones de contorno en las ecuaciones anotadas en los puntos $x_\alpha = l_\alpha + h^-_\alpha$ y $x_\alpha = L_\alpha - h^+_\alpha$. Estas ecuaciones se transforman, se hacen bipuntuales, y ellas es posible considerar como nuevas condiciones de contorno para ecuaciones tripuntuales de coeficientes constantes anotados para $l_\alpha + h^-_\alpha + h \leqslant x_\alpha \leqslant L_\alpha - h^+_\alpha - h$. Por consiguiente, tendremos los siguientes problemas para $v(x)$ y $w(x)$ (el índice superior α en v y w es omitido provisionalmente):

$$v^{+1\alpha} - 2v + v^{-1\alpha} = -1, \quad l_\alpha + h^-_\alpha + h \leqslant x_\alpha \leqslant L_\alpha - h^+_\alpha - h,$$

$$\left(1 + \frac{h}{h^-_\alpha}\right) v = v^{+1\alpha} + 1, \quad x_\alpha = l_\alpha + h^-_\alpha. \qquad (16)$$

$$\left(1+\frac{h_\alpha^+}{h}\right)v=\frac{h_\alpha^+}{h}v^{-1_\alpha}+1,\quad x_\alpha=L_\alpha-h_\alpha^+,$$

$$w^{+1_\alpha}-2w+w^{-1_\alpha}=0,\quad l_\alpha+h_\alpha^-+h\leqslant x_\alpha\leqslant L_\alpha-h_\alpha^+-h,$$

$$\left(1+\frac{h}{h_\alpha^-}\right)w=w^{+1_\alpha}-\frac{1}{2}\left(1-\frac{h}{h_\alpha^-}\right),\quad x_\alpha=l_\alpha+h_\alpha^-,$$

$$\left(1+\frac{h_\alpha^+}{h}\right)w=\frac{h_\alpha^+}{h}w^{-1_\alpha}+\frac{1}{2}\left(1-\frac{h_\alpha^+}{h}\right),\quad x_\alpha=L_\alpha-h_\alpha^+.$$

Utilizando los métodos de resolución de ecuaciones en diferencias de coeficientes constantes, expuestos en el § 4 del cap. I, hallaremos la forma explícita de la solución de los problemas de contorno (16) y (17):

$$v_\alpha(x)=\Big[\frac{1}{2h^2}(x_\alpha-l_\alpha)\Big(L_\alpha-x_\alpha+$$

$$+\frac{2h^2-(h_\alpha^++h_\alpha^-)(h_\alpha^++h-h_\alpha^-)}{L_\alpha-l_\alpha}\Big)+h_\alpha^-(h-h_\alpha^-)\Big],$$

$$w^\alpha(x)=\frac{1}{2}-\frac{h_\alpha^-(L_\alpha-x_\alpha)+h_\alpha^+(x_\alpha-l_\alpha)}{2h(L_\alpha-l_\alpha)}$$

para $l_\alpha+h_\alpha^-\leqslant x_\alpha\leqslant L_\alpha-h_\alpha^+$. Ya que $h_\alpha^\pm\leqslant h$, entonces

$$(h_\alpha^++h_\alpha^-)(h_\alpha^++h-h_\alpha^-)\geqslant h_\alpha^-(h-h_\alpha^-),$$

por eso

$$v^\alpha(x)\leqslant\frac{1}{2h^2}(x_\alpha-l_\alpha)(L_\alpha-x_\alpha)+1\leqslant\frac{1}{2h^2}\left(\frac{L_\alpha-l_\alpha}{2}\right)^2+1,$$

$$w^\alpha(x)\leqslant\frac{1}{2},\quad \alpha=1,\ 2.$$

Por consiguiente,

$$b_\alpha(x_\beta)=\max_{x_\alpha\in\omega_\alpha}v^\alpha(x)\leqslant\frac{1}{2h^2}\left(\frac{L_\alpha-l_\alpha}{2}\right)^2+1,$$

$$c_\alpha(x_\beta)=\max_{x_\alpha\in\omega_\alpha}\omega_\alpha(x)\leqslant\frac{1}{2}.$$

Por lo tanto, $\Delta=O(l_0^2/h^2)$, donde l_0 es el diámetro de la región G. Por eso en virtud del teorema 1 para el número de iteraciones es cierta la estimación

$$n\geqslant n_0(\varepsilon)=\frac{\ln(2/\varepsilon)}{2\sqrt{2}\sqrt[4]{\eta}}=\frac{\ln(2/\varepsilon)}{2\sqrt{2}\sqrt[4]{2}\sqrt{h/l_0}}\approx$$

$$\approx 0{,}298\sqrt{N}\ln\frac{2}{\varepsilon},\qquad (18)$$

donde N es el número máximo de nodos por la dirección x_1 o x_2. De esta forma, el número de iteraciones para el ejemplo modelo examinado depende solamente del paso fundamental h del retículo y no depende de los pasos en los nodos fronterizos de la red $\bar{\omega}$.

Comparemos la estimación (18) con la estimación del número de iteraciones para el caso del problema de Dirichlet en un cuadrado con lado l_0 y con un número N de nodos por cada dirección para la red cuadrada $\bar{\omega}$. La estimación correspondiente para el número de iteraciones fue obtenida en el punto 4 del § 1, y tiene la forma

$$n > n_0(\varepsilon) = 0{,}28\sqrt{N}\ln(2/\varepsilon).$$

De aquí se desprende que para una región arbitraria G el número de iteraciones del método alternativo triangular modificado es prácticamente el mismo que el número de iteraciones para el problema de Dirichlet para la ecuación de Poisson en un cuadrado, cuyo lado es igual al diámetro de la región G.

OBSERVACIÓN 1. Es evidente que el procedimiento expuesto aquí para construir el método alternativo triangular se puede utilizar también para el caso, cuando es necesario resolver una ecuación elíptica en un rectángulo, pero sobre una red no uniforme.

OBSERVACIÓN 2. La construcción del método para el caso de una ecuación con derivadas mixtas se puede llevar a efecto con ayuda de la elección de un regularizador, de forma semejante a como fue hecho en el punto 5 del § 2.

Capítulo XI

Método de las direcciones variables

En este capítulo se examinan métodos iterativos especiales para resolver ecuaciones elípticas reticulares $Au = f$, en las cuales el operador A posee una estructura determinada. En el § 1 es estudiado el método de las direcciones variables para el caso conmutativo; es construida la colección óptima de los parámetros. En el § 2 se ilustra el método en ejemplos de resolución de problemas de contorno para las ecuaciones elípticas de variables separables. El § 3 está dedicado al método de las direcciones variables en el caso no conmutativo.

§ 1. Método de las direcciones variables en el caso conmutativo

1. Esquema iterativo del método. En el capítulo X fue estudiado el método iterativo universal alternativo triangular, en el cual el operador B se elegía tomando en consideración la descomposición del operador A en suma de dos operadores conjugados uno a otro. Con más frecuencia se utiliza la descomposición de A en suma de operadores triangulares, siendo B en este caso un producto de operadores triangulares dependientes de un parámetro iterativo complementario. La toma en cuenta de la estructura del operador B permite elegir óptimamente los parámetros iterativos y construir un método, el cual converge esencialmente más rápido que el método explícito. En las aplicaciones a la resolución de ecuaciones elípticas reticulares este método es también económico, ya que en la realización de un paso iterativo se exige un número de operaciones aritméticas, proporcional al número de incógnitas en el problema.

Como sabemos, los operadores A correspondientes a las ecuaciones elípticas reticulares poseen una estructura espe-

cífica. Por eso durante la elección de los operadores B en los esquemas iterativos implícitos es natural intentar utilizar esta particularidad del operador A. Es evidente que tales métodos iterativos no serán universales, sin embargo, la restricción de la clase de los problemas iniciales por la exigencia de una determinada estructura del operador A permite construir métodos iterativos rápidamente convergentes, orientados precisamente a la resolución de ecuaciones reticulares.

En el presente capítulo será estudiado un método especial llamado *método iterativo de las direcciones variables*. Primeramente se da la descripción de este método en la forma operacional, y después en ejemplos será puesta en evidencia la aplicación de este método para encontrar una solución aproximada de diferentes ecuaciones elípticas reticulares.

Comenzaremos la descripción del método por el esquema iterativo. Supongamos que se exige hallar la solución de la ecuación operacional lineal

$$Au = f \tag{1}$$

con el operador no degenerado A definido en el espacio de Hilbert H. Sea el operador A representado en forma de la suma de los operadores A_1 y A_2, es decir, $A = A_1 + A_2$. Para la resolución aproximada de la ecuación (1) examinaremos el esquema iterativo implícito de dos capas:

$$B_{k+1}\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \quad y_0 \in H, \tag{2}$$

$$B_k = (\omega_k^1 E + A_1)(\omega_k^2 E + A_2), \quad \tau_k = \omega_k^{(1)} + \omega_k^{(2)}, \tag{3}$$

en el cual el operador B_{k+1} sobre la capa superior depende del número k de iteraciones. Aquí $\omega_k^{(1)}$ y $\omega_k^{(2)}$ son parámetros iterativos los cuales también dependen del número de iteraciones k y están sujetos a definición.

Detengámonos primeramente en los procedimientos para encontrar y_{k+1} siendo y_k dado. Uno de los posibles algoritmos de realización del esquema (2) es el siguiente:

$$\begin{aligned}(\omega_{k+1}^{(1)}E + A_1)\, y_{k+1/2} &= (\omega_{k+1}^{(1)}E - A_2)\, y_k + f,\\ (\omega_{k+1}^{(2)}E + A_2)\, y_{k+1} &= (\omega_{k+1}^{(2)}E - A_1)\, y_{k+1/2} + f,\\ & \qquad k = 0, 1, \ldots,\end{aligned} \tag{4}$$

donde $y_{k+1/2}$ es una aproximación iterativa intermedia.

Mostremos que el sistema (4) es algebraicamente equivalente al esquema (2). Para esto eliminaremos $y_{k+1/2}$ de (4). Teniendo en cuenta, que $A = A_1 + A_2$, volvamos a escribir (4) en la forma siguiente:

$$\begin{aligned}&(\omega_{k+1}^{(1)}E + A_1)(y_{k+1/2} - y_k) + Ay_k = f,\\ &(\omega_{k+1}^{(2)}E + A_2)(y_{k+1} - y_k) - \\ &\qquad - (\omega_{k+1}^{(2)}E - A_1)(y_{k+1/2} - y_k) + Ay_k = f\end{aligned} \tag{5}$$

y de la primera igualdad restemos la segunda. Obtendremos

$$y_{k+1/2} - y_k = (\omega_{k+1}^{(2)}E + A_2)\frac{y_{k+1} - y_k}{\omega_{k+1}^{(1)} + \omega_{k+1}^{(2)}} =$$

$$= (\omega_{k+1}^{(2)}E + A_2)\frac{y_{k+1} - y_k}{\tau_{k+1}}.$$

Sustituyendo esta expresión en (5), obtendremos el esquema (2). La transición inversa es evidente.

Para encontrar y_{k+1} se puede también utilizar otro algoritmo, interpretando (2) como un esquema con corrección w_k,

$$\begin{aligned}&(\omega_{k+1}^{(1)}E + A_1)\,v = r_k, \quad r_k = Ay_k - f,\\ &(\omega_{k+1}^{(2)}E + A_2)\,w_k = v,\\ &y_{k+1} = y_k - \tau_{k+1}w_k, \quad k = 0, 1, \ldots\end{aligned}$$

Este algoritmo es más económico en comparación con (4), sin embargo, exige memorizar una mayor información intermedia, es decir, exige memoria complementaria de la CE, lo que no es siempre cómodo.

Señalemos que como durante la construcción del esquema iterativo (3), así también al construir los algoritmos no impusimos ninguna condición sobre los operadores A_1 y A_2, excepto la suposición natural sobre la no degeneración de los operadores $\omega_k^{\alpha}E + A_\alpha$, $\alpha = 1, 2$. Todas las exigencias complementarias sobre los operadores A_1 y A_2 están relacionadas con el problema sobre la elección óptima de los parámetros $\omega_k^{(1)}$ y $\omega_k^{(2)}$.

2. Planteamiento del problema sobre la elección de los parámetros. En el método de las direcciones variables nosotros tenemos que ver con dos sucesiones de parámetros $\{\omega_k^{(1)}\}$ y $\{\omega_k^{(2)}\}$, los cuales eligiremos de la condición de mínimo de la norma del operador de resolución en el espacio inicial H.

Para resolver el problema sobre la elección de los parámetros iterativos es necesario hacer determinadas suposiciones respecto a los operadores A_1 y A_2, los cuales tienen un carácter funcional, y además prefijar alguna información a priori. Formulemos estas suposiciones.

Supondremos que el operador A se puede representar en forma de la suma de dos operadores autoconjugados y conmutables A_1 y A_2:

$$A = A_1 + A_2, \quad A_1 = A_1^*, \quad A_2 = A_2^*,$$
$$A_1 A_2 = A_2 A_1. \tag{6}$$

Sea la información a priori dada en forma de las cotas δ_α y Δ_α del operador A_α, $\alpha = 1, 2$, o sea,

$$\delta_1 E \leqslant A_1 \leqslant \Delta_1 E, \quad \delta_2 E \leqslant A_2 \leqslant \Delta_2 E, \tag{7}$$

y al mismo tiempo se cumple la condición

$$\delta_1 + \delta_2 > 0. \tag{8}$$

Notemos que de (6) — (8) se deduce la autoconjugación y definición positiva del operador A.

Si la suposición sobre la conmutatividad de los operadores A_1 y A_2 se cumple, entonces diremos que se examina el *caso conmutativo*, o de otro modo el *caso general*. Las condiciones (6) aseguran la autoconjugación de los operadores B_k para cualquier k. En efecto, en virtud de (6) los operadores $\omega_k^{(1)} E + A_1$ y $A_2 + \omega_k^{(2)} E$ son autoconjugados y conmutan, y el producto de operadores autoconjugados y conmutables es un operador autoconjugado.

Pasemos al estudio de la convergencia del esquema iterativo (2). Sustituyendo en (2) $y_k = z_k + u$, donde z_k es el error y u, la solución de la ecuación (1), obtendremos para z_k la ecuación homogénea

$$z_{k+1} = S_{k+1} z_k, \quad k = 0, 1, \ldots, \quad z_0 = y_0 - u, \tag{9}$$

donde

$$S_k = E - \tau_k B_k^{-1} A = (\omega_k^{(2)} E + A_2)^{-1} (\omega_k^{(1)} E + A_1)^{-1} \times \\ \times (\omega_k^{(2)} E - A_1)(\omega_k^{(1)} E - A_2). \tag{10}$$

Utilizando (9), expresemos z_n mediante z_0. Obtendremos

$$z_n = T_{n,0} z_0, \quad T_{n,0} = \prod_{j=1}^{n} S_j = S_n S_{n-1} \ldots S_1, \tag{11}$$

donde $T_{n,0}$ es el operador de resolución. Puesto que los operadores A_1 y A_2 son conmutables, entonces el orden de los factores en (10) es indiferente, todos los operadores S_k son autoconjugados y conmutan por parejas y, por consiguiente, el operador $T_{n,0}$ es autoconjugado en H: $T_{n,0} = R_n(A_1, A_2)$, donde $R_n(x, y)$ es un producto de funciones racionales fraccionales de x e y:

$$R_n(x, y) = \prod_{j=1}^{n} \frac{\omega_j^{(2)} - x}{\omega_j^{(1)} + x} \frac{\omega_j^{(1)} - y}{\omega_j^{(2)} + y}. \tag{12}$$

De (11) obtendremos

$$\| z_n \| \leqslant \| T_{n,0} \| \| z_0 \|. \tag{13}$$

En virtud de la autoconjugación del operador $T_{n,0}$ tenemos $\| T_{n,0} \| = \max\limits_k | \lambda_k (T_{n,0})$, donde $\lambda_k (T_{n,0})$ son los valores propios del operador $T_{n,0}$. A continuación, en virtud de las condiciones (6) (véase el punto 5 del § 1 del cap. V) los operadores A_1, A_2 y $T_{n,0}$ poseen un sistema común de funciones propias. Por eso

$$\lambda_k (T_{n,0}) = R_n (\lambda_{k_1}^{(1)}, \lambda_{k_2}^{(2)}),$$

donde $\lambda_{k_1}^{(1)}$ y $\lambda_{k_2}^{(2)}$ son los valores propios de los operadores A_1 y A_2 respectivamente, al mismo tiempo en virtud de (7) tenemos $\delta_1 \leqslant \lambda_{k_1}^{(1)} \leqslant \Delta_1$, $\delta_2 \leqslant \lambda_{k_2}^{(2)} \leqslant \Delta_2$. Por lo tanto,

$$\| T_{n,0} \| = \max_{k_1, k_2} | R_n (\lambda_{k_1}^{(1)}, \lambda_{k_2}^{(2)}) | \leqslant \max_{\substack{\delta_1 \leqslant x \leqslant \Delta_1 \\ \delta_2 \leqslant y \leqslant \Delta_2}} | R_n (x, y) |.$$

Sustituyendo estas estimaciones en (13), obtenemos

$$\| z_n \|_D \leqslant \max_{\substack{\delta_1 \leqslant x \leqslant \Delta_1 \\ \delta_2 \leqslant y \leqslant \Delta_2}} | R_n (x, y) | \| z_0 \|_D, \tag{14}$$

donde $R_n(x, y)$ está definido en (12), y $D = E$. Observemos que en virtud de la conmutabilidad de los operadores A_1 y A_2 el operador $T_{n,0}$ será autoconjugado también en el espacio energético H_D para $D = A, A_2$. Por eso, en virtud del lema 5 del § 1 del cap. V tendremos $\| T_{n,0} \| = \| T_{n,0} \|_A = \| T_{n,0} \|_{A^2}$ y, por consiguiente, la estimación (14) es válida para $D = A$ y $D = A^2$.

Así, el problema de estimar el error del esquema iterativo (2) está reducido al problema de estimar el máximo del módulo de una función de dos variables $R_n(x, y)$ en el rectán-

gulo $G = \{\delta_1 \leqslant x \leqslant \Delta_1,\ \delta_2 \leqslant y \leqslant \Delta_2\}$ y a la elección de los parámetros iterativos de la condición de mínimo del máximo del módulo de esta función. El problema planteado es bastante complejo y en el punto 3 se reducirá al problema más simple de encontrar la función racional fraccional de una variable que menos se desvía del cero sobre un segmento.

3. Transformación fracional lineal. Estudiemos la función $R(x, y)$. Por una transformación fracional lineal de las incógnitas aplicaremos el rectángulo G en el cuadrado $\{\eta \leqslant u \leqslant 1,\ \eta \leqslant v \leqslant 1,\ \eta > 0\}$, además eligiremos la transformación de tal forma, que ella no cambie el tipo de la función $R_n(x, y)$. La transformación buscada tiene la forma

$$x = \frac{ru - s}{1 - tu},\quad y = \frac{rv + s}{1 + tv},\quad \eta \leqslant u,\quad v \leqslant 1, \tag{15}$$

donde las constantes r, s, t y η están sujetas a definición.

Sustituyendo (15) en (12) e introduciendo nuevos parámetros $\varkappa_j^{(1)}$ y $\varkappa_j^{(2)}$,

$$\varkappa_j^{(1)} = \frac{\omega_j^{(1)} - s}{r - t\omega_j^{(1)}},\quad \varkappa_j^{(2)} = \frac{\omega_j^{(2)} + s}{r + t\omega_j^{(2)}},\quad j = 1, 2, \ldots, n, \tag{16}$$

obtenemos

$$R_n(x, y) = P_n(u, v) = \prod_{j=1}^{n} \frac{\varkappa_j^{(2)} - u}{\varkappa_j^{(1)} + u} \frac{\varkappa_j^{(1)} - v}{\varkappa_j^{(2)} + v}.$$

De (16) hallaremos las relaciones con ayuda de las cuales los parámetros $\omega_j^{(1)}$ y $\omega_j^{(2)}$ se expresan mediante los parámetros introducidos $\varkappa_j^{(1)}$ y $\varkappa_j^{(2)}$:

$$\omega_j^{(1)} = \frac{r\varkappa_j^{(1)} + s}{1 + t\varkappa_j^{(1)}},\quad \omega_j^{(2)} = \frac{r\varkappa_j^{(2)} - s}{1 - t\varkappa_j^{(2)}},\quad j = 1, 2, \ldots, n. \tag{17}$$

De esta forma, si son hallados los parámetros $\varkappa_j^{(1)}$ y $\varkappa_j^{(2)}$, por las fórmulas (17) se determinan los parámetros $\omega_j^{(1)}$ y $\omega_j^{(2)}$.

En virtud del cambio (15) llegamos al problema de buscar aquellos valores de los parámetros $\varkappa_j^{(1)}$ y $\varkappa_j^{(2)}$, para los cuales se obtiene

$$\min_{\varkappa^{(1)},\ \varkappa^{(2)}} \max_{\eta \leqslant u,\ v \leqslant 1} |P_n(u, v)|.$$

Notemos que si imponemos algunas limitaciones sobre la elección de los parámetros $\varkappa_j^{(1)}$ y $\varkappa_j^{(2)}$, por ejemplo $\varkappa_j^{(1)} = \varkappa_j^{(2)} = \varkappa_j$, entonces,

es evidente, que el mínimo puede sólo aumentar. Por lo tanto

$$\min_{\varkappa^{(1)},\,\varkappa^{(2)}} \max_{\eta\leqslant u,v\leqslant 1} |P_n(u, v)| \leqslant$$

$$\leqslant \min_{\varkappa} \max_{\eta\leqslant u,\,v\leqslant 1} \left| \prod_{j=1}^{n} \frac{\varkappa_j - u}{\varkappa_j + u} \frac{\varkappa_j - v}{\varkappa_j + v} \right| = \min_{\varkappa} \max_{\eta\leqslant u\leqslant 1} |r_n(u, \varkappa)|^2,$$

$$r_n(u, \varkappa) = \prod_{j=1}^{n} \frac{\varkappa_j - u}{\varkappa_j + u}.$$

Así pues, el problema planteado más arriba sobre la elección óptima de los parámetros iterativos $\omega_j^{(1)}$ y $\omega_j^{(2)}$ se ha reducido a encontrar la función racional fraccional $r_n(u, \varkappa)$ que menos se desvía del cero en el intervalo $[\eta, 1]$. De otro modo, es necesario hallar aquellos $\varkappa_j^{(*)}$, para los cuales

$$\max_{\eta\leqslant u\leqslant 1} |r_n(u, \varkappa^*) = \min_{\varkappa} \max_{\eta\leqslant u\leqslant 1} |r_n(u, \varkappa)| = \rho.$$

Si estos parámetros son hallados, para el error z_n de (14) se deducirá la estimación $\| z_n \|_D \leqslant \rho^2 \| z_0 \|_D$, y la exactitud ε será obtenida, si ponemos $\rho^2 = \varepsilon$.

La elección buscada de los parámetros iterativos será dada en el punto 4, y aquí hallaremos las constantes r, s, t y η de la transformación (15).

Si $r \neq ts$, entonces la transformación (15) es monótona según u y v, y, por consiguiente, la transformación inversa $u = (x + s)/(r + tx)$, $v = (y - s)/(r - ty)$ será monótona por x e y. Por eso, para aplicar el rectángulo $\{\delta_1 \leqslant x \leqslant \Delta_1,\ \delta_2 \leqslant y \leqslant \Delta_2\}$ sobre el cuadrado $\{\eta \leqslant u, v \leqslant 1\}$ es suficiente, que los extremos del segmento $[\delta_\alpha, \Delta_\alpha]$ pasen a los extremos del segmento $[\eta, 1]$. Esto nos da cuatro relaciones para determinar las constantes de la transformación (15):

$$\delta_1 = \frac{r\eta - s}{1 - t\eta},\quad \delta_2 = \frac{r\eta + s}{1 + t\eta},\quad \Delta_1 = \frac{r - s}{1 - t},\quad \Delta_2 = \frac{r + s}{1 + t}. \tag{18}$$

Hallemos la solución del sistema no lineal (18). Notemos primeramente, que en virtud de la suposición (8) son válidas las desigualdades

$$\Delta_2 + \delta_1 \geqslant \delta_1 + \delta_2 > 0,\quad \Delta_1 + \delta_2 \geqslant \delta_2 + \delta_1 > 0. \tag{19}$$

Después, de (18) obtenemos

$$\begin{aligned} \Delta_1 - \delta_1 &= \frac{(1-\eta)(r-st)}{(1-t)(1-t\eta)},\quad \Delta_2 - \delta_2 = \frac{(1-\eta)(r-st)}{(1+t)(1+t\eta)}, \\ \Delta_2 + \delta_1 &= \frac{(1+\eta)(r-st)}{(1+t)(1-t\eta)},\quad \Delta_1 + \delta_2 = \frac{(1+\eta)(r-st)}{(1-t)(1+t\eta)}. \end{aligned} \tag{20}$$

De aquí hallaremos

$$\left(\frac{1-\eta}{1+\eta}\right)^2 = \frac{(\Delta_1 - \delta_1)(\Delta_2 - \delta_2)}{(\Delta_1 + \delta_2)(\Delta_2 + \delta_1)} < 1,$$

y como en virtud de (19) el denominador no se anula, entonces

$$\eta = \frac{1-a}{1+a}, \quad a = \sqrt{\frac{(\Delta_1-\delta_1)(\Delta_2-\delta_2)}{(\Delta_1+\delta_2)(\Delta_2+\delta_1)}}, \quad \eta \in [0, 1]. \tag{21}$$

Hallemos ahora t. De (20) obtenemos

$$\frac{\Delta_2+\delta_1}{\Delta_1-\delta_1} = \frac{1+\eta}{1-\eta}\,\frac{1-t}{1+t} = \frac{1}{a}\,\frac{1-t}{1+t}.$$

De aquí tendremos

$$t = \frac{1-b}{1+b}, \quad b = \frac{\Delta_2+\delta_1}{\Delta_1-\delta_1}\,a. \tag{22}$$

De las dos últimas ecuaciones del sistema (18) hallaremos

$$r = \frac{1}{2}\,[\Delta_1(1-t)+\Delta_2(1+t)] = \frac{1+t}{2}\,[\Delta_2+\Delta_1 b] = \frac{\Delta_2+\Delta_1 b}{1+b}, \tag{23}$$

$$s = \frac{1}{2}\,[\Delta_2(1+t)-\Delta_1(1+t)] = \frac{1+t}{2}\,[\Delta_2-\Delta_1 b] = \frac{\Delta_2-\Delta_1 b}{1+b}. \tag{24}$$

Ya que

$$r-st = \frac{2b(\Delta_1+\Delta_2)}{(1+b)^2} > 0, \quad |t|<1,$$

entonces la transformación (15) efectivamente es monótona. En el punto 4 será mostrado que $\eta < \varkappa_j \equiv \varkappa_j^{(1)} \equiv \varkappa_j^{(2)} < 1$, por eso en (17) el denominador no se anula.

Examinemos algunos ejemplos. Sean $\delta_1 = \delta_2 = \delta$ y $\Delta_1 = \Delta_2 = \Delta$, es decir, las cotas de los operadores A_1 y A_2 son iguales. Entonces $\eta = \delta/\Delta$, $t = s = 0$, $r = \Delta$, $w_j^{(1)} \equiv w_j^{(2)} = \Delta\varkappa_j$. Sean ahora $\delta_1 = 0$, $\delta_2 = \delta$, $\Delta_1 = \Delta_2 = \Delta$, es decir, el operador A_1 es degenerado. Entonces

$$\eta = \delta/(\Delta+\sqrt{\Delta^2-\delta^2}), \quad t = \eta, \quad s = \Delta\eta, \quad r = \Delta,$$

$$\omega_j^{(1)} = \frac{\Delta\varkappa_j+\Delta\eta}{1+\eta\varkappa_j}; \quad \omega_j^{(2)} = \frac{\Delta\varkappa_j-\Delta\eta}{1-\eta\varkappa_j}, \quad j = 1, 2, \ldots, n.$$

4. **Conjunto óptimo de parámetros.** Citemos la resolución del problema sobre la elección óptima de los parámetros iterativos. A diferencia del caso cuando encontramos el polinomio con menor desviación de cero, el cual fue examinado en el § 2 del cap. VI, aquí los parámetros iterativos $\varkappa_j$ no se expresan mediante funciones trigonométricas, sino con ayuda de las funciones elípticas de Jacobi.

Recordemos algunas definiciones. La integral definida

$$K(k) = \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1-k^2\,\mathrm{sen}^2\,\varphi}}$$

se llama *integral elíptica completa de primer género*, el número k se llama *módulo* de esta integral, y el número $k' = \sqrt{1-k^2}$ se llama *módulo complementario*. Se acostumbra designar $K(k') = K'(k)$.

Si designamos mediante $u(z, k)$ la función

$$u(z, k) = \int_z^1 \frac{dy}{\sqrt{(1-y^2)(y^2-(k^2))}},$$

entonces la función $z = \mathrm{dn}(u, k')$, inversa a $u(z, k)$, se llama *función elíptica de Jacobi* de argumento u y módulo k'.

Utilizando estas designaciones, la solución exacta del problema sobre la elección óptima de los parámetros iterativos integrales $\varkappa_j$ se puede escribir en la forma siguiente:

$$\varkappa_j \in \mathfrak{M}_n = \left\{\mu_i = \mathrm{dn}\left(\frac{2i-1}{2n} K'(\eta), \eta'\right), \quad i = 1, 2, \ldots, n\right\}, \tag{25}$$

$j = 1, 2, \ldots, n$,

donde el número n de iteraciones, suficiente para alcanzar una exactitud ε, se estima por la fórmula

$$n \geqslant u_0(\varepsilon) = \frac{1}{4} \frac{K'(\eta)}{K(\eta)} \frac{K'(\varepsilon)}{K(\varepsilon)}. \tag{26}$$

Aquí, como en el método de Chebishev, en calidad de $\varkappa_j$ se eligirán consecutivamente todos los elementos del conjunto $\mathfrak{M}_n$. Formularemos los resultados obtenidos para el método de las direcciones variables en el caso conmutativo en forma de teorema.

TEOREMA 1. *Sean cumplidas las condiciones* (6) — (8), *y los parámetros* $\omega_j^{(1)}$ *y* $\omega_j^{(2)}$ *elegidos por las fórmulas*

$$\omega_j^{(1)} = \frac{r\varkappa_j + s}{1 + t\varkappa_j}, \quad \omega_j^{(2)} = \frac{r\varkappa_j - s}{1 - t\varkappa_j}, \quad j = 1, 2, \ldots, n, \tag{27}$$

donde $\varkappa_j$ *y* n *están definidos en* (25), (26) *y* r, s, t, η en (21) — (24). *El método de las direcciones variables* (2), (3) *converge en* H_D, *y después de ser cumplimentadas* n *iteracciones para el error* $z_n = y_n - u$ *será cierta la estimación* $\| z_n \|_D \leqslant \varepsilon \| z_0 \|_D$, *donde* $D = E$, A *o* A^2, *y* n *se determina de acuerdo con* (26).

Volvamos ahora a la parte computacional de la realización del método de las direcciones variables con una lista óptima de parámetros. Hallemos las fórmulas aproximadas para calcular $\varkappa_j$ y n e indiquemos el orden, en el cual se deben elegir los parámetros del conjunto $\mathfrak{M}_n$.

Utilizando la representación asintótica de las integrales elípticas completas para pequeños valores de k:

$$\frac{1}{K(k)} = \frac{2}{\pi} + O(k^2), \quad K'(k) = \ln\frac{4}{k} + O\left(k^2 \ln\frac{1}{k}\right),$$

de (26) obtendremos la siguiente fórmula aproximada para el número de iteraciones n:

$$n \geqslant n_0(\varepsilon) = \frac{1}{\pi^2} \ln\frac{4}{\eta} \ln\frac{4}{\varepsilon}. \tag{28}$$

Examinemos ahora la pregunta sobre el cálculo de μ_i. La función dn (u, k') decrece monótonamente respecto de u, tomando los siguientes valores: dn $(0, k') = 1$, dn $(K'(k), (k') = K$. Por eso $\eta < \mu_n < \mu_{n-1} < \ldots < \mu_1 < 1$. A continuación, de las propiedades de la función elíptica dn (u, k'):

$$dn(u, k') = \frac{k}{dn(K'(k) - u, k')}$$

se deduce que tiene lugar la igualdad

$$\mu_i = \eta/\mu_{n+1-i}, \quad i = 1, 2, \ldots \tag{29}$$

Por eso es suficiente hallar la mitad de los valores de μ_i, y los demás determinar de las relaciones (29).

Obtendremos una fórmula aproximada para μ_i, utilizando el desarrollo de la función dn (u, k') en potencias de k. Para eso expresamos la función dn $(\sigma K'(\eta), \eta')$ mediante la theta-función de Jacobi, y representemos estas funciones por series. Obtendremos

$$dn(\sigma K'(\eta), \eta') = \frac{\sqrt{\eta}\,\theta_3\left(\frac{i\sigma\pi K'}{K}, \bar{q}\right)}{\theta_3\left(\frac{i\sigma\pi K'}{K'}, \bar{q}\right)} =$$

$$= \sqrt{\eta}\,\bar{q}^{\frac{2\sigma-1}{4}} \frac{\sum\limits_{m=-\infty}^{\infty} \bar{q}^{m(m+\sigma)}}{\sum\limits_{m=-\infty}^{\infty} \bar{q}^{m(m-1+\sigma)}},$$

donde $\bar{q} = \exp\left(-\frac{\pi K'(\eta)}{K(\eta)}\right) = \frac{\eta^2}{16}\left(1 + \frac{\eta^2}{2}\right) + O(\eta^6)$.

De aquí encontramos

$$dn(\sigma K'(\eta), \eta') = \sqrt{\eta}\, q^{\frac{2\sigma-1}{4}} \frac{1 + q^{1-\sigma} + q^{1+\sigma}}{1 + q^{\sigma} + q^{2-\sigma}} + O(\eta^{\nu}); \tag{30}$$

donde

$$q = \eta^2(1 + \eta^2/2)/16, \quad \nu = \begin{cases} 4 + 5\sigma, & 0 < \sigma < 1/2, \\ 8 - 3\sigma, & 1/2 \leqslant \sigma < 1. \end{cases}$$

Para $\sigma \geqslant 1/2$ el orden del término residual en (30) es uniformemente igual a 5 respecto a σ, y para $\sigma < 1/2$ el orden es igual a 4. Por eso la fórmula aproximada para $dn\,(\sigma K'(\eta), \eta')$ será más exacta para $\sigma \geqslant 1/2$, que para $\sigma < 1/2$.

De (25), (29) y (30) obtenemos las siguientes fórmulas para calcular μ_l:

$$\mu_l = \sqrt{\eta}\, q^{\frac{2\sigma_l - 1}{4}} \frac{1 + q^{1-\sigma_l} + q^{1+\sigma_l}}{1 + q^{\sigma_l} + q^{2-\sigma_l}}, \quad [n/2] + 1 \leqslant l \leqslant n,$$

$$\mu_l = \eta/\mu_{n+1-l}, \quad 1 \leqslant i \leqslant [n/2], \quad \sigma_i = (2i - 1)/(2n),$$

$$q = \eta^2 (1 + \eta^2/6)/16,$$

donde $[a]$ es la parte entera de a.

Examinemos ahora la pregunta sobre el orden de elección de $\varkappa_j$ del conjunto $\mathfrak{M}_n$. De la definición del operador de transición S_j en el esquema (2) y de las propiedades (6) y (7), obtenemos

$$\| S_j \| = \max_k |\lambda_k (S_j)| \leqslant \max_{\substack{\delta_1 \leqslant x \leqslant \Delta_1 \\ \delta_2 \leqslant y \leqslant \Delta_2}} \left| \frac{\omega_j^{(2)} - x}{\omega_j^{(1)} + x} \frac{\omega_j^{(1)} - y}{\omega_j^{(2)} + y} \right|$$

o en virtud del cambio (15)

$$\| S_j \| \leqslant \max_{\eta \leqslant u \leqslant 1} \left| \frac{\varkappa_j - u}{\varkappa_j + u} \right|^2.$$

Ya que todos los $\varkappa_j$ pertenecen al intervalo $(\eta, 1)$, entonces de aquí se deduce que $\| S_j \| < 1$ para cualquier j. Por eso el método iterativo (2), (3) será estable a los errores de redondeo para cualquier orden de elección de los $\varkappa_j$ del conjunto $\mathfrak{M}_n$, por ejemplo, $\varkappa_j = \mu_j$, $j = 1, 2, \ldots, n$.

Como conclusión de este punto mostremos que para la lista (conunto) construida de los parámetros $\omega_j^{(1)}$ y $\omega_j^{(2)}$ los operadores $\omega_j^{\alpha} E + A_\alpha$, $\alpha = 1, 2$, son definidos positivos en H para cualquier j. En efecto, de (27) obtenemos

$$\frac{\partial \omega_j^{(1)}}{\partial \varkappa_j} = \frac{r - st}{(1 + t\varkappa_j)^2} > 0, \quad \frac{\partial \omega_j^{(2)}}{\partial \varkappa_j} = \frac{r - st}{(1 - t\varkappa_j)^2} > 0.$$

Puesto que los denominadores en (27) no se anulan y $\eta < \varkappa_j < 1$, entonces de aquí y de (18) hallaremos

$$\delta_2 = \frac{r\eta + s}{1 + t\eta} \leqslant \omega_j^{(1)} \leqslant \frac{r + s}{1 + t} = \Delta_2; \quad \delta_1 = \frac{r\eta - s}{1 - t\eta} \leqslant \omega_j^{(2)} \leqslant$$

$$\leqslant \frac{r - s}{1 - t} = \Delta_1. \qquad (31)$$

Por consiguiente, en virtud de la suposición (7) de (31) obtendremos

$$\omega_j^{(1)} E + A_1 \geqslant (\delta_1 + \delta_2) E, \quad \omega_j^{(2)} E + A_2 \geqslant (\delta_1 + \delta_2) E,$$

y como $\delta_1 + \delta_2 > 0$ por la suposición (8), entonces la afirmación está demostrada.

§ 2. Ejemplos de aplicación del método

1. Problemas de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo. Comenzaremos el examen de los ejemplos de aplicación del método de las direcciones variables por la resolución del problema de Dirichlet de diferencias para la ecuación de Poisson en un rectángulo.

Supongamos que sobre la red rectangular $\bar{\omega} = \{x_{ij} = (ih_1,\ jh_2) \in \bar{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1, 2\}$, introducida en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$, se exige hallar la solución del problema

$$\Lambda y = (\Lambda_1 + \Lambda_2)\, y = -\varphi(x),\quad x \in \omega,\quad y(x) = g(x),\ x \in \gamma,$$
$$\Lambda_\alpha y = y_{\bar{x}_\alpha x_\alpha},\quad \alpha = 1, 2. \tag{1}$$

Designemos mediante H el espacio de las funciones reticulares definidas sobre ω con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2.$$

Definiremos los operadores A, A_1 y A_2 sobre H de la forma siguiente: $Ay = -\Lambda \overset{\circ}{y}$, $A_\alpha y = -\Lambda_\alpha \overset{\circ}{y}$, $\alpha = 1, 2$, donde $y \in H$, $\overset{\circ}{y} \in \overset{\circ}{H}$, $y(x) = \overset{\circ}{y}(x)$ para $x \in \omega$, y $\overset{\circ}{H}$ es el conjunto de las funciones reticulares definidas sobre $\bar{\omega}$ y que se anulan sobre γ.

El problema de diferencias (1) puede entonces ser escrito en forma de la ecuación operacional $Au = f$, donde $A = A_1 + A_2$.

Como nosotros sabemos (véase el punto 5 del § 1 del cap. V), los operadores A_α son autoconjugados en H y poseen cotas δ_α y Δ_α:

$$\delta_\alpha = \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha},\quad \Delta_\alpha = \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2l_\alpha},\quad \alpha = 1, 2,$$

las cuales coinciden con los valores propios mínimo y máximo de los operadores de diferencias Λ_α. Queda por verificar el cumplimiento de la condición de conmutabilidad de los operadores A_1 y A_2. Utilizando la definición de los operadores A_α y de los operadores de diferencias Δ_α, obtendremos

$$A_1 A_2 y = \overset{\circ}{y}_{\bar{x}_1 x_1 \bar{x}_2 x_2} = \overset{\circ}{y}_{\bar{x}_2 x_2 \bar{x}_1 x_1} = A_2 A_1 y,$$

lo que se exigía demostrar.

Así, las condiciones requeridas para aplicar el método de las direcciones variables en el caso conmutativo, para el ejemplo a examinar están cumplidas.

Utilizando la definición de los operadores A_1 y A_2, el algoritmo del método de las direcciones variables para el ejemplo a examinar se puede escribir en la forma siguiente:

$$\omega_{k+1}^{(1)} y_{k+1/2} - \Lambda_1 y_{k+1/2} = \omega_{k+1}^{(1)} y_k + \Lambda_2 y_k + \varphi,$$
$$h_1 \leqslant x_1 \leqslant l_1 - h_1, \quad (2)$$
$$y_{k+1/2}(x) = g(x), \quad x_1 = 0, \ l_1, \quad h_2 \leqslant x_2 \leqslant l_2 - h_2,$$
$$\omega_{k+1}^{(2)} y_{k+1} - \Lambda_2 y_{k+1} = \omega_{k+1}^{(2)} y_{k+1/2} + \Lambda_1 y_{k+1/2} + \varphi,$$
$$h_2 \leqslant x_2 \leqslant l_2 - h_2, \quad (3)$$
$$y_{k+1} = g(x), \quad x_2 = 0, \ l_2, \quad h_1 \leqslant x_1 \leqslant l_1 - h_1,$$

al mismo tiempo $y_k(x) = g(x)$ siendo $x \in \gamma$ para cualquier $k \geqslant 0$. De esta forma, el algoritmo del método consiste en la resolución sucesiva para cada x_2 fijo de los problemas de contorno tripuntuales (2) por la dirección x_1 para determinar $y_{k+1/2}$ sobre ω y en la resolución para cada x_1 de los problemas de contorno (3) por la dirección x_2 para determinar una nueva aproximación iterativa y_{k+1} sobre ω. La alternación de las direcciones, por las cuales se resuelven los problemas de contorno (2), (3) dio el nombre del método: método de las direcciones variables.

Para encontrar la solución de los problemas (2), (3) se puede aprovechar el método de factorización. Escribamos las ecuaciones (2), (3) en forma de un sistema tripuntual y verifiquemos el cumplimiento de las restantes condiciones de estabilidad del método de factorización. Las ecuaciones tendrán la forma

$$-y_{k+1/2}(i+1, j) + (2 + h_1^2 \omega_{k+1}^{(1)}) y_{k+1/2}(i, j) -$$
$$- y_{k+1/2}(i-1, j) = \varphi_1(i, j), \quad 1 \leqslant i \leqslant N_1 - 1, \quad (4)$$
$$y_{k+1/2}(0, j) = g(0, j), \quad y_{k+1/2}(N_1, j) = g(N_1, j),$$
$$1 \leqslant j \leqslant N_2 - 1,$$

donde

$$\varphi_1(i, j) = \frac{h_1^2}{h_2^2} [y_k(i, j+1) - (2 + h_2^2 \omega_{k+1}^{(1)}) y_k(i, j) +$$
$$+ y_k(i, j-1) + h_2^2 \varphi(i, j)];$$
$$-y_{k+1}(i, j+1) + (2 + h_2^2 \omega_{k+1}^{(2)}) y_{k+1}(i, j) - y_{k+1}(i, j-1) =$$
$$= \varphi_2(i, j), \quad (5)$$
$$1 \leqslant j \leqslant N_2 - 1,$$

$y_{k+1}(i, 0) = g(i, 0), \quad y_{k+1}(i, N_2) = g(i, N_2),$

$1 \leqslant i \leqslant N_1 - 1,$

donde

$$\varphi_2(i, j) = \frac{h_2^2}{h_1^2}[y_{k+1/2}(i+1, j) -$$

$$- (2 + h_1^2 \omega_{k+1}^{(2)}) y_{k+1/2}(i, j) + y_{k+1/2}(i-1, j) + h_2^2 \varphi(i, j)]$$

Puesto que para el ejemplo dado $\delta_1 > 0$ y $\delta_2 > 0$, entonces en virtud de las desigualdades (31) del punto 4 del § 1 los parámetros $\omega_k^{(1)}$ y $\omega_k^{(2)}$ son positivos. Por lo tanto, en las ecuaciones tripuntuales (4) y (5) los coeficientes de $y_{k+1/2}(i, j)$ y de $y_{k+1}(i, j)$ predominan sobre los restantes coeficientes. Por consiguiente, el método de factorización aplicado a los problemas (4) y (5) será estable respecto a los errores de redondeo.

Calculemos el número de operaciones aritméticas que es necesario efectuar para realizar un paso iterativo en el método (2), (3) para el ejemplo a examinar. Es suficiente contar el número de operaciones para el problema (4), el cálculo para el problema (5) se realiza de modo análogo.

Las fórmulas del método de factorización para el problema (4) tienen la forma (j está fijo):

$$y_{k+1/2}(i, j) = \alpha_i y_{k+1/2}(i+1, j) + \beta_i, \quad 1 \leqslant i \leqslant N_1 - 1,$$

$$y_{k+1/2}(N_1, j) = g(N_1, j),$$

$$\alpha_{i+1} = 1/(C - \alpha_i), \quad i = 1, 2, \ldots, N_1 - 1,$$

$$\alpha_1 = 0, \quad C = 2 + h_1^2 \omega_{k+1}^{(1)},$$

$$\beta_{i+1} = \alpha_{i+1}(\varphi_1(i, j) + \beta_i),$$

$$i = 1, 2, \ldots, N_1 - 1, \quad \beta_1 = g(0, j).$$

Notemos que los coeficientes de factorización α_i no dependen de j y por eso pueden ser calculados una vez con un gasto de $2(N_1 - 1)$ operaciones aritméticas. Después, en el cálculo de $\varphi_1(i, j)$ sobre la red ω se exige $6(N_1 - 1) \times (N_2 - 1)$ operaciones aritméticas. Los coeficientes de factorización β_i y la solución $y_{k+1/2}$ es necesario calcularlos de nuevo para cada j. Para esto se exige $4(N_1 - 1) \times (N_2 - 1)$ operaciones. En total para encontrar $y_{k+1/2}$ sobre la red ω siendo dada y_k se exige $Q_1 = 10(N_1 - 1) \times (N_2 - 1) + 2(N_1 - 1)$ operaciones aritméticas. Para

encontrar y_{k+1} de (15) por la $y_{k+1/2}$ calculada se exige $Q_2 = 10\,(N_1 - 1)\,(N_2 - 1) + 2\,(N_2 - 1)$ operaciones. Así pues, para el ejemplo a examinar la realización de un paso iterativo en el método de las direcciones variables se efectúa por

$$Q = 20\,(N_1 - 1)\,(N_2 - 1) + 2\,(N_1 - 1) + 2\,(N_2 - 1) \quad (6)$$

operaciones aritméticas.

Estimemos ahora el número de iteraciones n suficiente para obtener la solución con una exactitud ε prefijada. En el caso particular, cuando la región $\overline{G}$ es un cuadrado con lado l ($l_1 = l_2 = l$) y la red $\overline{\omega}$ es cuadrada con $N_1 = N_2 = N$ ($h_1 = h_2 = l/N$), tendremos

$$\delta_1 = \delta_2 = \delta = \frac{4}{h^2} \operatorname{sen}^2 \frac{\pi h}{2l}, \quad \Delta_1 = \Delta_2 = \Delta = \frac{4}{h^2} \cos^2 \frac{\pi h}{2l}.$$

De (21) y (28) del § 1 obtenemos la siguiente estimación para el número de iteraciones

$$n \geqslant n_0(\varepsilon) = 0{,}1 \ln \frac{4}{\eta} \ln \frac{4}{\varepsilon}, \quad \eta = \delta/\Delta = \operatorname{tg}^2 \frac{\pi h}{2l}$$

ó para los h pequeños

$$n_0(\varepsilon) = 0{,}2 \ln (4N/\pi) \ln (4/\varepsilon), \quad (7)$$

es decir, el número de iteraciones es proporcional al logaritmo del número de incógnitas N por una dirección.

De (6) y (7) obtendremos la siguiente estimación para el número de operaciones aritméticas $Q(\varepsilon)$ gastadas en encontrar la solución del problema de diferencias (1) por el método de las direcciones variables con una exactitud ε:

$$Q(\varepsilon) = nQ = 4N^2 \ln (4N/\pi) \ln (4/\varepsilon). \quad (8)$$

Para comparar este método con el método directo de reducción completa (véase el § 3 del cap. III), pasaremos en (8) de los logaritmos naturales a los logaritmos de base 2. Obtendremos

$$Q(\varepsilon) \approx 2{,}12N^2 \log_2 (4N/\pi) \log_2 (4/\varepsilon).$$

Ya que el error de la aproximación de esquema de diferencias (1) es $O(h^2)$, conviene elegir ε igual a $O(h^2)$.

Si tomamos $\varepsilon = 4/N^2$, entonces obtendremos

$$Q(\varepsilon) = 4{,}24N^2 \log_2 N \log_2 (4N/\pi).$$

Para $N = 64$ obtendremos $\varepsilon \approx 10^{-3}$ y

$$Q(\varepsilon) \approx 27{,}6N^2 \log_2 N.$$

La comparación con la estimación del número de operaciones para el método de reducción completa muestra que para la red indicada el método de las direcciones variables exige aproximadamente en 5,5 veces más operaciones aritméticas que el método de reducción completa. Con el aumento de N y la disminución de ε esta diferencia aumenta.

Para el caso particular a examinar citemos el número de iteraciones n en dependencia del número de nodos N por una dirección para $\varepsilon = 10^{-4}$.

Para comparar citemos el número de iteraciones para otros métodos examinados más arriba.

Tabla 11

N	Método de iteración simple	Método de Chebishev explícito	Método de relajación superior	Método alternativo triangular	Método de las direcciones variables
32	1 909	101	65	16	8
64	7 642	202	128	23	10
128	30 577	404	257	32	11

De la tabla se desprende que el menor número de iteraciones se exige para el método de las direcciones variables. Por el número de iteraciones ante él cede el método alternativo triangular con los parámetros de Chebishev, el cual fue examinado en el capítulo X.

OBSERVACIÓN. Si para el problema a examinar (1) analizamos el método de las direcciones variables con parámetros constantes, es decir $\omega_j^1 \equiv \omega^{(1)}$, $\omega_j^{(2)} \equiv \omega^{(2)}$, $\tau_j = \omega^{(1)} + \omega^{(2)}$, entonces de la fórmula (25) del § 1 obtenemos en virtud de la igualdad $\operatorname{dn}\left(\frac{1}{2}K'(k), k'\right) = \sqrt{k}$, que el parámetro $\varkappa_j \equiv \sqrt{\eta}$. En el caso particular, cuando $\delta_1 = \delta_2 = \delta$, $\Delta_1 = \Delta_2 = \Delta$, con anterioridad en el punto 3 del § 1 fue obtenida la siguiente relación entre los parámetros $\omega_j^{(1)}$, $\omega_j^{(2)}$ y $\varkappa_j$: $\omega_j^{(1)} = \omega_j^{(2)} = \Delta\varkappa_j$. Ya que en este caso $\eta = \delta/\Delta$, entonces de aquí encontramos $\omega^{(1)} = \omega^{(2)} = \sqrt{\delta\Delta}$.

2. **Tercer problema de contorno para una ecuación elíptica con variables separables.** Supongamos que en el rectán-

gulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \alpha = 1, 2\}$ se exige hallar la solución del siguiente problema de contorno:

$$Lu = \sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha}\left(k_\alpha(x_\alpha)\frac{\partial u}{\partial x_\alpha}\right) - qu = -f(x), \quad x \in G,$$

$$k_\alpha(x_\alpha)\frac{\partial u}{\partial x_\alpha} = \varkappa_{-\alpha}u - g_{-\alpha}(x), \quad x_\alpha = 0, \tag{9}$$

$$-k_\alpha(x_\alpha)\frac{\partial u}{\partial x_\alpha} = \varkappa_{+\alpha}(u) - g_{+\alpha}(x), \quad x_\alpha = l_\alpha, \quad \alpha = 1, 2.$$

Supondremos que se cumplen las siguientes condiciones:

$$0 < C_{1,\alpha} \leqslant k_\alpha(x_\alpha) \leqslant c_{2,\alpha}, \ \varkappa_{\pm\alpha} = \text{const} \geqslant 0, \tag{10}$$
$$\alpha = 1, 2, \quad q = \text{const} \geqslant 0, \ \sum_{\alpha=1}^{2} \varkappa_{\pm\alpha}^2 + q^2 \neq 0.$$

El problema de contorno de Neumann ($\varkappa_{\pm\alpha} = 0$) para el caso $q = 0$ será examinado por separado en el capítulo XII. Las condiciones (10) aseguran la existencia y unicidad de la solución del problema (9).

Sobre la red rectangular $\overline{\omega} = \{x_{ij} = (ih_1, jh_2) \in \overline{G}, 0 \leqslant i \leqslant N_1, 0 \leqslant j \leqslant N_2, h_\alpha = l_\alpha/N_\alpha, \alpha = 1, 2\}$ al problema (9) le corresponde el problema de contorno de diferencias

$$\Lambda y = (\Lambda_1 + \Lambda_2)y = -\varphi(x), \quad x \in \overline{\omega}, \tag{11}$$

donde dos operadores de diferencias Λ_1 y Λ_2, y el segundo miembro φ se determinan de la manera siguiente:

$$\Lambda_\alpha y = \begin{cases} \frac{2}{h_\alpha} a_\alpha(h_\alpha) y_{x_\alpha} - \left(0{,}5q + \frac{2}{h_\alpha}\varkappa_{-\alpha}\right) y, & x_\alpha = 0, \\ (a_\alpha(x_\alpha) y_{\bar{x}_\alpha})_{x_\alpha} - 0{,}5qy, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ -\frac{2}{h_\alpha} a_\alpha(l_\alpha) y_{\bar{x}_\alpha} - \left(0{,}5q + \frac{2}{h_\alpha}\varkappa_{+\alpha}\right) y, & x_\alpha = l_\alpha \end{cases}$$

para $0 \leqslant x_\beta \leqslant l_\beta$, $\beta = 3 - \alpha$; $\alpha = 1, 2$ y $\varphi = f + \varphi_1 + \varphi_2$,

$$\varphi_\alpha(x) = \begin{cases} \frac{2}{h_\alpha} g_{-\alpha}(x), & x_\alpha = 0, \\ 0, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ \frac{2}{h_\alpha} g_{+\alpha}(x), & x_\alpha = l_\alpha. \end{cases}$$

Designemos mediante H el espacio de las funciones reticulares definidas sobre $\overline{\omega}$, en el cual el producto escalar está definido por la fórmula

$$(u, v) = \sum_{x \in \overline{\omega}} u(x) v(x) \hbar_1(x_1) \hbar_2(x_2), \quad x \in \overline{\omega},$$

$$\hbar_\alpha(x_\alpha) = \begin{cases} 0{,}5h_\alpha, & x_\alpha = 0, \\ h_\alpha, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha. \end{cases}$$

Definiremos los operadores A, A_1 y A_2 sobre H por las relaciones $Ay = -\Lambda y$, $A_\alpha y = -\Lambda_\alpha y$, $\alpha = 1, 2$. En el § 2 del cap. V fue mostrado que los operadores A_1 y A_2 así definidos son autoconjugados y conmutables. Además, en virtud de las condiciones (10) el operador A es definido positivo en H (es decir $\delta_1 + \delta_2 > 0$). Quedan por hallar las cotas de los operadores A_1 y A_2, es decir, las constantes δ_α y Δ_α en las desigualdades $\delta_\alpha E \leqslant A_\alpha \leqslant \Delta_\alpha E$, $\alpha = 1, 2$.

Hallemos primeramente δ_α. De la definición de los operadores A_α y de las fórmulas de Green de diferencias obtenemos

$$(A_\alpha y, y) = -\sum_{x_\beta=0}^{l_\beta} \sum_{x_\alpha=h_\alpha}^{l_\alpha - h_\alpha} [(a_\alpha y_{\bar{x}_\alpha})_{x_\alpha} - 0{,}5qy]\, y \hbar_1 \hbar_2 -$$

$$- \sum_{x_\beta=0}^{l_\beta} \left[a_\alpha(h_\alpha) y_{x_\alpha} - \left(\varkappa_{-\alpha} + \frac{h_1}{4} q \right) y \right] y \,|_{x_\alpha=0} \hbar_2 +$$

$$+ \sum_{x_\beta=0}^{l_\beta} \left[a_\alpha(l_\alpha) y_{\bar{x}_\alpha} + \left(\varkappa_{+\alpha} + \frac{h_1}{4} q \right) y \right] y \,|_{x_\alpha=l_\alpha} \hbar_2 =$$

$$= \sum_{x_\beta=0}^{l_\beta} \sum_{x_\alpha=h_\alpha}^{l_\alpha} a_\alpha y_{\bar{x}_\alpha}^2 \hbar_1 \hbar_2 +$$

$$+ \sum_{x_\beta=0}^{l_\beta} (\varkappa_{-\alpha} y^2 |_{x_\alpha=0} + \varkappa_{+\alpha} y^2 |_{x_\alpha=l_\alpha}) \hbar^2 + 0{,}5q(y^2, 1).$$

De aquí hallaremos, que si $q = \varkappa_{-\alpha} = \varkappa_{+\alpha} = 0$, entonces $\delta_\alpha = 0$. Si al menos una de las magnitudes q, $\varkappa_{-\alpha}$ ó $\varkappa_{+\alpha}$ es diferente de cero, entonces δ_α se puede hallar de la

siguiente forma. En virtud del lema 16 del § 2 del cap. V tendremos

$$(y^2, 1)_{\bar{\omega}_\alpha} \leqslant \max_{x_\alpha \in \bar{\omega}_\alpha} v^\alpha(x_\alpha)\,(A_\alpha y, y)_{\bar{\omega}_\alpha}, \tag{12}$$

donde $v^\alpha(x_\alpha)$ es la solución del problema de contorno tripuntual

$$\begin{aligned}
&(a_\alpha(x_\alpha)\, v^\alpha_{\bar{x}_\alpha})_{x_\alpha} - 0{,}5qv = -1, \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,\\
&\frac{2}{h_\alpha} a_\alpha(h_\alpha)\, v^\alpha_{x_\alpha} - \left(0{,}5q + \frac{2}{h_\alpha}\varkappa_{-\alpha}\right) v^\alpha = -1, \quad x_\alpha = 0,\\
&-\frac{2}{h_\alpha} a_\alpha(l_\alpha)\, v^\alpha_{\bar{x}_\alpha} - \left(0{,}5q + \frac{2}{h_\alpha}\varkappa_{+\alpha}\right) v^\alpha = -1, \quad x_\alpha = l_\alpha,
\end{aligned} \tag{13}$$

y el producto escalar se define de la manera siguiente:

$$(u, v)_{\bar{\omega}_\alpha} = \sum_{x_\alpha = 0}^{l_\alpha} u(x)\, v(x)\, \hbar_\alpha(x_\alpha).$$

Multiplicando (12) por $\hbar_\beta(x_\beta)$ y adicionando por x_β desde 0 hasta l_β, obtendremos

$$(y^2, 1) \leqslant \max_{x_\alpha \in \bar{\omega}_\alpha} v^\alpha(x_\alpha)\,(A_\alpha y, y)$$

y, por consiguiente

$$\delta_\alpha = \frac{1}{\max\limits_{x_\alpha \in \bar{\omega}_\alpha} v^\alpha(x_\alpha)}, \quad \alpha = 1, 2.$$

Hallemos ahora Δ_α. Al operador A_α le corresponde una matriz tridiagonal $\mathscr{A}_\alpha$. Designemos mediante $\mathscr{D}$ la parte diagonal de la matriz A_α es decir, $\mathscr{D}y = d_\alpha(x_\alpha)\, y$,

$$d_\alpha(x_\alpha) = \begin{cases}
0{,}5q + \dfrac{2}{h_\alpha}\varkappa_{-\alpha} + \dfrac{2}{h_\alpha^2} a_\alpha(h_\alpha), & x_\alpha = 0,\\
0{,}5q + \dfrac{1}{h_\alpha^2}\left(a_\alpha(x_\alpha) + a_\alpha(x_\alpha) + h_\alpha\right), & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,\\
0{,}5q + \dfrac{2}{h_\alpha}\varkappa_{+\alpha} + \dfrac{2}{h_\alpha^2} a_\alpha(l_\alpha), & x_\alpha = l_\alpha.
\end{cases}$$

Examinemos el problema de valores propios

$$A_\alpha y - \lambda \mathscr{D} y = 0, \quad x \in \bar{\omega}. \tag{14}$$

Es fácil mostrar que si λ es un valor propio del problema (14), entonces $2 - \lambda$ también es un valor propio. Por lo tanto,

$$\lambda_{\min} \mathscr{D} \leqslant A_\alpha \leqslant (2 - \lambda_{\min})\, \mathscr{D}$$

ó

$$(A_\alpha y, y) \leqslant (2 - \lambda_{\min})\, (\mathscr{D} y, y) \leqslant (2 - \lambda_{\min}) \max_{x_\alpha \in \bar{\omega}_\alpha} d_\alpha (x_\alpha)\, (y, y).$$

Por esto en calidad de Δ_α se puede tomar

$$\Delta_\alpha = (2 - \lambda_{\min}) \max_{x_\alpha \in \bar{\omega}_\alpha} d_\alpha (x_\alpha).$$

Queda por hallar $\lambda_{\min}$. Si $q = \varkappa_{-\alpha} = \varkappa_{+\alpha} = 0$, entonces el operador A_α es degenerado y $\lambda_{\min} = 0$. De otro modo, en virtud de la observación 2 del lema 14 del § 2 del cap. V, tendremos

$$(d_\alpha y, y)_{\bar{\omega}_\alpha} \leqslant \max_{\bar{x}_\alpha \in \bar{\omega}_\alpha} w^\alpha (x_\alpha)\, (A_\alpha y, y)_{\bar{\omega}_\alpha}, \tag{15}$$

donde $w^\alpha (x_\alpha)$ es la solución del siguiente problema de contorno:

$$\begin{aligned}
&(a_\alpha w^\alpha_{\bar{x}_\alpha})_{x_\alpha} - 0{,}5 q w^\alpha = -d_\alpha (x_\alpha), \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,\\
&\frac{2}{h_\alpha} a_\alpha (h_\alpha)\, w^\alpha_{x_\alpha} - \left(0{,}5q + \frac{2}{h_\alpha} \varkappa_{-\alpha}\right) w^\alpha = -d_\alpha (0), \quad x_\alpha = 0,\\
&-\frac{2}{h_\alpha} a_\alpha (l_\alpha)\, w^\alpha_{\bar{x}_\alpha} - \left(0{,}5q + \frac{2}{h_\alpha} \varkappa_{+\alpha}\right) w^\alpha = -d_\alpha (l_\alpha),\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad x_\alpha = l_\alpha.
\end{aligned} \tag{16}$$

Multiplicando (15) por $h_\beta (x_\beta)$ y sumando según x_β desde 0 hasta l_β obtendremos

$$(\mathscr{D} y, y) \leqslant \max_{\bar{x}_\alpha \in \bar{\omega}_\alpha} w^\alpha (x_\alpha)\, (A_\alpha y, y)$$

y por, consiguiente,

$$\lambda_{\min} \geqslant \frac{1}{\max\limits_{x_\alpha \in \bar{\omega}_\alpha} w^\alpha (x_\alpha)}.$$

De esta forma, si $q = \varkappa_{-\alpha} = \varkappa_{+\alpha} = 0$, entonces

$$\delta_\alpha = 0, \quad \Delta_\alpha = 2 \max_{x_\alpha \in \bar{\omega}_\alpha} d_\alpha (x_\alpha),$$

ó de otro modo

$$\delta_\alpha = \frac{1}{\max\limits_{x_\alpha \in \bar{\omega}_\alpha} v^\alpha(x_\alpha)},$$

$$\Delta_\alpha = \left(2 - \frac{1}{\max\limits_{x_\alpha \in \bar{\omega}_\alpha} w^\alpha(x_\alpha)}\right) \max_{x_\alpha \in \bar{\omega}_\alpha} d_\alpha(x_\alpha),$$

donde $v^\alpha(x_\alpha)$ y $w^\alpha(x_\alpha)$ son las soluciones de los problemas (13) y (16). Toda la información a priori necesaria para aplicar el método de las direcciones variables está hallada. Utilizando las fórmulas del teorema 1, hallaremos los parámetros iterativos del método y estimaremos el número de iteraciones exigido.

Citemos ahora las fórmulas del algoritmo del método de las direcciones variables para el ejemplo a examinar. Teniendo en cuenta la definición de los operadores A_1, A_2 y del segundo miembro de φ, obtendremos

$$\omega^{(1)}_{k+1} y_{k+1/2} - \Lambda_1 y_{k+1/2} = \omega^{(2)}_{k+1} y_k + \Lambda_2 y_k + \varphi,$$
$$0 \leqslant x_1 \leqslant l_1, \quad 0 \leqslant x_2 \leqslant l_2,$$
$$\omega^{(2)}_{k+1} y_{k+1} - \Lambda_2 y_{k+1} = \omega^{(2)}_{k+1} y_{k+1/2} + \Lambda_1 y_{k+1/2} + \varphi,$$
$$0 \leqslant x_2 \leqslant l_2, \quad 0 \leqslant x_1 \leqslant l_1.$$

Aquí, a diferencia del problema de Dirichlet, los problemas de contorno tripuntuales deben resolverse también por la frontera de la red $\bar{\omega}$, y la aproximación inicial y_0 es una función reticular arbitraria definida sobre toda la red $\bar{\omega}$.

Utilizando la condición (10), se puede mostrar que para el ejemplo a examinar, como para el caso del problema de Dirichlet, para el número de iteraciones n es válida la siguiente estimación asintótica respecto a h:

$$n \geqslant n_0(\varepsilon) = O(\ln |h| \ln \varepsilon), \quad |h|^2 = h_1^2 + h_2^2.$$

Señalemos que todos los análisis aquí realizados conservan su validez también en el caso, cuando $\bar{\omega}$ es una red rectangular no uniforme arbitraria en la región $\bar{G}$. Unicamente conviene cambiar los operadores Λ_α aquí introducidos por operadores sobre una red no uniforme.

Subrayemos que las suposiciones de la constancia de q, $\varkappa_{\pm\alpha}$ y de la dependencia de los coeficientes a_α sólo de x_α son esenciales. Si no se cumple al menos una de estas suposi-

ciones, entonces la condición de conmutatividad de los operadores A_1 y A_2 no será satisfecha.

Como conclusión señalemos que el método de las direcciones variables puede aplicarse para resolver análogos de diferencias de la ecuación (9) y bajo otras condiciones de contorno. En particular, sobre cada lado del rectángulo $\overline{G}$ puede ser dada una de las condiciones de contorno de primero, segundo o tercer género con los constantes $\varkappa_{\pm\alpha}$.

3. **Problemas de Dirichlet de diferencias de orden de exactitud elevado.** Examinemos otro ejemplo de aplicación del método de las direcciones variables. Supongamos que sobre la red rectangular $\overline{\omega} = \{(x_{ij} = (ih_1, jh_2) \in \overline{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1, 2\}$, introducida en el rectángulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$, se exige hallar la solución del problema de Dirichlet de diferencias de orden de exactitud elevado para la ecuación de Poisson

$$\Lambda y = (\Lambda_1 + \Lambda_2 + (\varkappa_1 + \varkappa_2)\,\Lambda_1\Lambda_2)\,y = -\varphi(x), \quad x \in \omega, \qquad (17)$$

$$y(x) = g(x), \quad x \in \gamma,$$

donde $\Lambda_\alpha y = y_{\bar{x}_\alpha x_\alpha}$, $\varkappa_\alpha = h_\alpha^2/12$, $\alpha = 1,\ 2$.

Aquí

$$\varphi = \widetilde{f} + \varkappa_1\Lambda_1\widetilde{f} + \varkappa_2\Lambda_2\widetilde{f},$$

donde $\widetilde{f}(x)$ es el segundo miembro de la ecuación diferencial inicial

$$Lu = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -\widetilde{f}(x), \quad x \in G, \quad u(x) = g(x), \quad x \in \Gamma.$$

El esquema de diferencias (17) para la elección indicada, de $\varphi(x)$ tiene una exactitud $O(|h|^4)$, $|h|^2 = h_1^2 + h_2^2$, y sobre una red cuadrada ($h_1 = h_2 = h$) para la correspondiente elección de $\varphi(x)$

$$\varphi = \widetilde{f} + \frac{h^2}{12}(\Lambda_1 + \Lambda_2)\widetilde{f} + \frac{h^4}{360}(\Lambda_1^2 + 4\Lambda_1\Lambda_2 + \Lambda_2^2)\widetilde{f}$$

tiene exactitud $O(h^6)$.

Introduciendo los operadores $A_\alpha y = -\Lambda_\alpha \mathring{y}$, donde $y \in H$, $\mathring{y} \in \mathring{H}$ y H es el espacio de las funciones reticulares, definidas sobre ω con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x) v(x) h_1 h_2,$$

y $\mathring{H}$ es el conjunto de las funciones reticulares que se anulan sobre γ, escribamos (17) en la forma operacional

$$Au = f, \tag{18}$$

donde $A = A_1 + A_2 - (\varkappa_1 + \varkappa_2) A_1 A_2$.

Como ha sido mostrado reiteradamente, los operadores A_1 y A_2 poseen las siguientes propiedades: A_1 y A_2 son autoconjugados en H y conmutables

$$A_\alpha = A_\alpha^*, \quad \alpha = 1, 2, \quad A_1 A_2 = A_2 A_1, \tag{19}$$

el operador A_α tiene cotas δ_α y Δ_α, donde

$$\delta_\alpha = \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2 l_\alpha}, \quad \Delta_\alpha = \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2 l_\alpha}, \tag{20}$$

$$\delta_\alpha E \leqslant A_\alpha \leqslant \Delta_\alpha E, \quad \delta_\alpha > 0, \quad \alpha = 1, 2.$$

Tiene lugar

LEMA 1. *Si están cumplidas las condiciones (19), (20) y* $\varkappa_\alpha \Delta_\alpha < 1$, *entonces los operadores*

$$\bar{A}_\alpha = (E - \varkappa_\alpha A_\alpha)^{-1} A_\alpha, \quad \alpha = 1, 2, \tag{21}$$

son autoconjugados en H, conmutables y tienen cotas $\bar{\delta}_\alpha$ *y* $\bar{\Delta}_\alpha$, *es decir*

$$\bar{\delta}_\alpha E \leqslant \bar{A}_\alpha \leqslant \bar{\Delta}_\alpha E, \quad \bar{\delta}_\alpha > 0, \quad \alpha = 1, 2,$$

donde $\bar{\delta}_\alpha$ y $\bar{\Delta}_\alpha$ *se determinan por las fórmulas*

$$\bar{\delta}_\alpha = \frac{\delta_\alpha}{1 - \varkappa_\alpha \delta_\alpha}, \quad \bar{\Delta}_\alpha = \frac{\Delta_\alpha}{1 - \varkappa_\alpha \Delta_\alpha}. \tag{22}$$

En efecto, la existencia del operador $\bar{A}_\alpha$ se desprende de la definición positiva del operador $E - \varkappa_\alpha A_\alpha$ si está satisfecha la condición $\varkappa_\alpha \Delta_\alpha < 1$. A continuación, representando $\bar{A}_\alpha$ en la forma $\bar{A}_\alpha = (A_\alpha^{-1} - \varkappa_\alpha E)^{-1}$, y teniendo en cuenta la autoconjugación de los operadores A_α, A_α^{-1} y $A_\alpha^{-1} - \varkappa_\alpha E$, obtenemos

$$\left(\frac{1}{\Delta_\alpha} - \varkappa_\alpha\right) E \leqslant (A_\alpha^{-1} - \varkappa_\alpha E) \leqslant \left(\frac{1}{\delta_\alpha} - \varkappa_\alpha\right) E.$$

De aquí se deduce la afirmación del lema. La conmutatividad de los operadores $\bar{A}_1$ y $\bar{A}_2$ se deduce de la conmutatividad de A_1 y A_2. El lema está demostrado.

Para el ejemplo a examinar las condiciones del lema 1, son cumplidas ya que $\varkappa_\alpha \Delta_\alpha < 1/3$.

Transformemos ahora la ecuación (18). Para esto escribamos (18) en la forma

$$A_1 (E - \varkappa_2 A_2)\, u + A_2 (E - \varkappa_1 A_1)\, u = f. \qquad (23)$$

Aplicando a (23) el operador $(E - \varkappa_1 A_1)^{-1} (E - \varkappa_2 A_2)^{-1}$ y teniendo en cuenta la conmutatividad de todos los operadores, obtendremos de (23) la ecuación equivalente a (18)

$$\bar{A}u = (\bar{A}_1 + \bar{A}_2)\, u = \bar{f}, \qquad (24)$$

donde $\bar{A}_1$ y $\bar{A}_2$ están definidos en (21), mientras $\bar{f} = (E - \varkappa_1 A_1)^{-1} (E - \varkappa_2 A_2)^{-1} f$. Así, se ha reducido la resolución de la ecuación (18) a la resolución de la ecuación (24) con los operadores autoconjugados y conmutables $\bar{A}_1$ y $\bar{A}_2$, cuyas cotas están prefijadas en (22).

Para encontrar una solución aproximada de la ecuación (24) utilizaremos el método de las direcciones variables

$$\bar{B}_{k+1} \frac{y_{k+1} - y_k}{\tau_{k+1}} + \bar{A} y_k = \bar{f}, \quad k = 0, 1, \ldots, y_0 \in H, \qquad (25)$$

donde

$$\bar{B}_k = (\omega_k^{(1)} E + \bar{A}_1)(\omega_k^{(2)} E + \bar{A}_2), \quad \tau_k = \omega_k^{(1)} + \omega_k^{(2)}.$$

Los parámetros iterativos $\omega_k^{(1)}$ y $\omega_k^{(2)}$ se encuentran por las fórmulas del teorema 1, en las cuales δ_α y Δ_α se sustituyen por $\bar{\delta}_\alpha$ y $\bar{\Delta}_\alpha$. Aquí se cumplen todas las condiciones necesarias para aplicar el método de las direcciones variables.

Queda por examinar el algoritmo que realiza el método iterativo (25). Volvamos a escribir (25) en la forma

$$(\omega_{k+1}^{(1)} E + \bar{A}_1)(\omega_{k+1}^{(2)} E + \bar{A}_2)\, y_{k+1} = \\ = (\omega_{k+1}^{(2)} E - \bar{A}_1)(\omega_{k+1}^{(1)} E - \bar{A}_2)\, y_k + \tau_{k+1} \bar{f}. \qquad (26)$$

En el punto 4 del § 1 fue mostrado que los parámetros iterativos $\omega_k^{(1)}$ y $\omega_k^{(2)}$ para cualquier k satisfacen las desigualdades $\bar{\delta}_2 \leqslant \omega_k^{(1)} \leqslant \bar{\Delta}_2$, $\bar{\delta}_1 \leqslant \omega_k^{(2)} \leqslant \bar{\Delta}_1$.

Ya que para el ejemplo a examinar $\bar{\delta}_\alpha > 0$, entonces, dividiendo los miembros primero y segundo de (26) por

$\omega_{k+1}^{(1)}\omega_{k+1}^{(2)}$ y designando $\tau_{k+1}^{(1)} = 1/\omega_{k+1}^{(1)}$, $\tau_{k+1}^{(2)} = 1/\omega_{k+1}^{(2)}$, obtendremos

$$(E + \tau_{k+1}^{(1)}\overline{A}_1)(E + \tau_{k+1}^{(2)}\overline{A}_2)\, y_{k+1} = \\ = (E - \tau_{k+1}^{(2)}\overline{A}_1)(E - \tau_{k+1}^{(1)}\overline{A}_2) + (\tau_{k+1}^{(1)} + \tau_{k+2}^{(2)})\,\overline{f}.$$

Apliquemos a ambos miembros de esta igualdad el operador

$$(E - \varkappa_1 A_1)(E - \varkappa_2 A_2)$$

y tomemos en cuenta que todos los operadores son conmutables y

$$(E + \varkappa_\alpha A_\alpha)(E + \tau_{k+1}^{\alpha}\overline{A}_\alpha) = E + (\tau_{k+1}^{\alpha} - \varkappa_\alpha)A_\alpha,$$
$$(E - \varkappa_\alpha A_\alpha)(E - \tau_{k+1}^{(\beta)}\overline{A}_\alpha) = E - (\tau_{k+1}^{(\beta)} + \varkappa_\alpha)A_\alpha,$$
$$\beta = 3 - \alpha, \quad \alpha = 1, 2.$$

Como resultado obtendremos

$$(E + (\tau_{k+1}^{(1)} - \varkappa_1)A_1)(E + (\tau_{k+1}^{(2)} - \varkappa_2)A_2)\, y_{k+2} = \\ = (E - (\tau_{k+1}^{(2)} + \varkappa_1)A_1)(E - (\tau_{k+1}^{(1)} + \varkappa_2)A_2)\, y_k + \\ + (\tau_{k+1}^{(1)} + \tau_{k+1}^{(1)})\, f. \quad (27)$$

El esquema iterativo (27) es equivalente al esquema siguiente

$$(E + (\tau_{k+1}^{(1)} - \varkappa_1)A_1)\, y_{k+1/2} = \\ = (E - (\tau_{k+1}^{(1)} + \varkappa_2)A_2)\, y_k + (\tau_{k+1}^{(1)} - \varkappa_1)\, f, \quad (28)$$
$$(E + (\tau_{k+1}^{(2)} - \varkappa_2)A_2)\, y_{k+1} = \\ = (E - (\tau_{k+1}^{(2)} + \varkappa_1)A_1)\, y_{k+1/2} + (\tau_{k+1}^{(2)} + \varkappa_1)\, f. \quad (29)$$

La equivalencia de (27) y (28), (29) se demuestra de la siguiente forma. Multiplicando (28) por $\tau_{k+1}^{(2)} + \varkappa_1$, (29) por $-(\tau_{k+1}^{(1)} - \varkappa_1)$ y adicionando los resultados, obtenemos

$$(\tau_{k+1}^{(1)} + \tau_{k+1}^{(2)})\, y_{k+1/2} = (\tau_{k+1}^{(1)} - \varkappa_1)(E + (\tau_{k+1}^{(2)} - \varkappa_2)A_2)\, y_{k+1} + \\ + (\tau_{k+1}^{(2)} + \varkappa_1)(E - (\tau_{k+1}^{(1)} + \varkappa_2)A_2)\, y_k. \quad (30)$$

Sustituyendo (30) en (28), obtendremos (27) después de transformaciones no complejas. El paso inverso de los razonamientos es evidente.

Teniendo en cuenta las definiciones de los operadores A_1 y A_2, el esquema (28) y (29) se puede escribir en la forma del esquema de diferencia usual

$$(E-(\tau_{k+1}^{(1)}-\varkappa_1)\Lambda_1)\,y_{k+1/2}=$$
$$=(E+(\tau_{k+1}^{(1)}+\varkappa_2)\Lambda_2)\,y_k+(\tau_{k+1}^{(1)}-\varkappa_1)\varphi \quad (31)$$

para $h_1\leqslant x_1\leqslant l_1-h_1$,

$$y_{k+1/2}=g(x)+(\varkappa_1+\varkappa_2)\Lambda_2 g(x), \quad x_1=0,\ l_1.$$

El problema de contorno (31) se debe resolver sucesivamente para $h_2\leqslant x_2\leqslant l_2-h_2$. Como resultado será hallado $y_{k+1/2}$ para $0\leqslant x_1\leqslant l_1$, $h_2\leqslant x_2\leqslant l_2-h_2$. Luego,

$$(E-(\tau_{k+1}^{2}-\varkappa_2)\Lambda_2)\,y_{k+1}=$$
$$=(E+(\tau_{k+1}^{(2)}+\varkappa_1)\Lambda_1)\,y_{k+1/2}+(\tau_{k+1}^{(2)}+\varkappa_1)\varphi \quad (32)$$

para $h_2\leqslant x_2\leqslant l_2-h_2$.

$$y_{k+1}=g(x), \quad x_2=0,\ l_2.$$

El problema de contorno (32) se debe resolver sucesivamente para $h_1\leqslant x_1\leqslant l_1-h_1$. Como resultado será hallado y_{k+1}.

Si comparamos el número de iteraciones del método de las direcciones variables para el esquema de diferencias de segundo orden de exactitud examinado en el punto 1 del § 2, y para el esquema con orden de exactitud elevado, entonces en el último caso el número de iteraciones será algo mayor. En el caso particular $l_1=l_2=l$, $N_1=N_2=N$ para $N=10$ este aumento ocurre en 1%, y para $N=100$ en 4%. El volumen de cálculo en cada iteración para ambos esquemas es prácticamente el mismo, y la diferencia en el número de iteraciones no es significativa. Puesto que el esquema con orden de exactitud elevado permite utilizar una red más burda para alcanzar una exactitud dada de la solución del problema diferencial, entonces su aplicación es especialmente ventajosa en aquellos casos, cuando la solución del problema diferencial posee suficiente suavidad.

Recordemos, que en el § 3 del cap. III para resolver el problema (17) nosotros examinamos un método directo: el método de reducción. Al igual que para un esquema de segundo orden de exactitud, para el esquema aquí a examinar el método directo exigirá un menor número de operaciones aritméticas que el método de las direcciones variables con parámetros óptimos.

§ 3. Método de las direcciones variables en el caso general

1. Caso de operadores no conmutativos. Supongamos que se exige hallar la solución de la ecuación operacional lineal

$$Au = f \tag{1}$$

con el operador no degenerado A, el cual representaremos en forma de la suma de dos operadores autoconjugados no conmutativos A_1 y A_2 con cotas δ_1, Δ_1 y δ_2, Δ_2:

$$A_\alpha = A_\alpha^*, \quad \delta_\alpha E \leqslant A_\alpha \leqslant \Delta_\alpha E, \quad \alpha = 1, 2, \quad \delta_1 + \delta_2 > 0, \tag{2}$$
$$A = A_1 + A_2.$$

Para la resolución aproximada de la ecuación (1) examinaremos el esquema de dos capas del método de las direcciones variables con dos parámetros iterativos $\omega^{(1)}$ y $\omega^{(2)}$:

$$\begin{aligned} &B\frac{y_{k+1}-y_k}{\tau} + Ay_k = f, \quad k = 0, 1, \ldots, \ y_0 \in H, \\ &B = (\omega^{(1)}E + A_1)(\omega^{(2)}E + A_2), \quad \tau = \omega^{(1)} + \omega^{(2)}. \end{aligned} \tag{3}$$

Aquí los parámetros iterativos y el operador B no dependen del número de la iteración k.

Como en el caso de los operadores conmutativos A_1 y A_2, la aproximación iterativa y_{k+1} para el esquema (3) puede ser hallada por el algoritmo siguiente:

$$(\omega^{(1)}E + A_1)\, y_{k+1/2} = (\omega^{(1)}E - A_2)\, y_k + f,$$
$$(\omega^{(2)}E + A_2)\, y_{k+1} = (\omega^{(2)}E - A_1)\, y_{k+1/2} + f, \quad k = 0, 1, \ldots.$$

Investiguemos la convergencia del esquema iterativo (3) y hallemos los valores óptimos de los parámetros $\omega^{(1)}$ y $\omega^{(2)}$. Suponiendo que el operador $\omega^{(2)}E + A_2$ no es degenerado, estudiaremos la convergencia de (3) en el espacio energético H_D, donde $D = (\omega^{(2)}E + A_2)^2$. En virtud de (2) el operador D es autoconjugado en H, y de la suposición indicada más arriba se sigue que D es definido positivo en H.

Para el error $z_k = y_k - u$ obtendremos de (3) la ecuación homogénea

$$z_{k+1} = Sz_k, \quad k = 0, 1, \ldots, \quad z_0 = y_0 - u,$$
$$S = (\omega^{(2)}E + A_2)^{-1}(\omega^{(1)}E + A_1)^{-1}(\omega^{(2)}E - A_1)(\omega^{(1)}E - A_2).$$

Pasemos en (4) al problema para el error equivalente

$$x_k = (\omega^{(2)}E + A_2)z_k. \tag{5}$$

Obtenemos

$$x_{k+1} = \bar{S} x_k, \quad k = 0, 1, \ldots, \quad \bar{S} = \bar{S}_1 \bar{S}_2,$$
$$\bar{S}_1 = (\omega^{(1)} E + A_1)^{-1} (\omega^{(2)} E - A_1), \tag{6}$$
$$\bar{S}_2 = (\omega^{(2)} E + A_2)^{-1} (\omega^{(1)} E - A_2).$$

Puesto que en virtud del cambio hecho (5) tiene lugar la igualdad $\| x_k \| = \| z_k \|_D$ entonces es suficiente investigar el comportamiento de la norma del error equivalente x_k en el espacio H. De (6) hallaremos

$$\| x_{k+1} \| \leqslant \| \bar{S} \| \| x_k \| \leqslant \| \bar{S}_1 \| \| \bar{S}_2 \| \| x_k \|, \quad k = 0, 1, \ldots$$

y, por consiguiente,

$$\| z_n \|_D = \| x_n \| \leqslant \| \bar{S} \|^n \| x_0 \| \leqslant (\| \bar{S}_1 \| \| \bar{S}_2 \|)^n \| z_0 \|_D. \tag{7}$$

Estimemos la norma de los operadores $\bar{S}_1$ y $\bar{S}_2$. Supongamos que los operadores $\omega_{(\alpha)} E + A_\alpha$, $\alpha = 1, 2$, son no negativos. Entonces del punto 4 del § 1 del cap. V en virtud de (2) obtendremos

$$\| \bar{S}_1 \| \leqslant \max_{\delta_1 \leqslant x \leqslant \Delta_1} \left| \frac{\omega^{(2)} - x}{\omega^{(1)} + x} \right|,$$
$$\| \bar{S}_2 \| \leqslant \max_{\delta_2 \leqslant y \leqslant \Delta_2} \left| \frac{\omega^{(1)} - y}{\omega^{(2)} + y} \right|$$

y, por lo tanto,

$$\| \bar{S}_1 \| \| \bar{S}_2 \| \leqslant \max_{\substack{\delta_1 \leqslant x \leqslant \Delta_1 \\ \delta_2 \leqslant y \leqslant \Delta_2}} | R_1 (x, y) |,$$
$$R_1 (x, y) = \frac{\omega^{(2)} - x}{\omega^{(1)} + x} \frac{\omega^{(1)} - y}{\omega^{(2)} + y}.$$

Tomando en consideración la estimación (7), planteemos el problema de elegir los parámetros $\omega^{(1)}$ y $\omega^{(2)}$ de la condición

$$\min_{\omega^{(1)}, \omega^{(2)}} \max_{\substack{\delta_1 \leqslant x \leqslant \Delta_1 \\ \delta_2 \leqslant y \leqslant \Delta_2}} | R_1 (x, y) |.$$

Este problema es un caso particular del problema resuelto en el § 1 del presente capítulo. Con ayuda de una transformación lineal fraccional de las variables x y v (vease (15), (21)—

(24) en el § 1) el problema planteado se reduce al problema de encontrar el parámetro $\varkappa^*$ de la condición

$$\max_{\eta \leqslant u \leqslant 1} \left| \frac{\varkappa^* - u}{\varkappa^* + u} \right| = \min_{\varkappa} \max_{\eta \leqslant u \leqslant 1} \left| \frac{\varkappa - u}{\varkappa + u} \right| = \rho. \qquad (8)$$

Con esto los parámetros $\omega^{(1)}$ y $\omega^{(2)}$ se expresan mediante $\varkappa^*$ por las fórmulas

$$\omega^{(1)} = \frac{r\varkappa^* + s}{1 + t\varkappa^*}, \quad \omega^{(2)} = \frac{r\varkappa^* - s}{1 - t\varkappa^*},$$

y para el error z_n tiene lugar la estimación

$$\| z_n \|_D \leqslant \rho^{2n} \| z_0 \|_D.$$

Además, en el punto 4 del § 1 fue mostrado que para la elección óptima de $\varkappa^*$ los operadores $\omega^\alpha E + A_\alpha$ son definidos positivos, si $\delta_1 + \delta_2 > 0$. Por consiguiente, en virtud de (2) nuestras suposiciones sobre la no negatividad de los operadores $\omega^\alpha E + A_\alpha$, $\alpha = 1, 2$, serán cumplidas a ciencia cierta.

Citemos la solución del problema (8) independientemente de los resultados del punto 4 del § 1. Examinemos la función $\varphi(u) = (\varkappa - u)/(\varkappa + u)$ sobre el segmento $0 \leqslant \eta \leqslant u \leqslant 1$ para $\varkappa > 0$. Esta función decrece monótonamente respecto de u y, por lo tanto,

$$\max_{\eta \leqslant u \leqslant 1} |\varphi(u)| = \max\left(\left| \frac{\eta - \varkappa}{\eta + \varkappa} \right|, \left| \frac{1 - \varkappa}{1 + \varkappa} \right| \right).$$

De aquí es fácil obtener, que el mínimo según $\varkappa$ de esta expresión se alcanza para $\varkappa^*$, el cual se determina de la igualdad

$$\frac{\varkappa^* - N}{\varkappa^* + N} = \frac{1 - \varkappa^*}{1 + \varkappa^*}.$$

De aquí hallaremos

$$\varkappa^* = \sqrt{\eta}, \quad \min_{\varkappa} \max_{\eta \leqslant u \leqslant 1} \left| \frac{\varkappa - u}{\varkappa + u} \right| = \rho = \frac{1 - \sqrt{\eta}}{1 + \sqrt{\eta}}.$$

Así, está demostrado

TEOREMA 2. *Sean cumplidas las condiciones* (2), *y los parámetros* $\omega^{(1)}$ *y* $\omega^{(2)}$ *elegidos por las fórmulas*

$$\omega^{(1)} = \frac{r\sqrt{\eta} + s}{1 + t\sqrt{\eta}}, \quad \omega^{(2)} = \frac{r\sqrt{\eta} - s}{1 - t\sqrt{\eta}},$$

donde r, s, t *y* η *están definidos en* (21)—(24) *del* § 1. *El método de las direcciones variables* (3) *converge en* H_D, *y para el error* z_n *tiene lugar la estimación*

$$\| z_n \|_D \leqslant \rho^{2n} \| z_0 \|_D, \quad \rho = \frac{1-\sqrt{\eta}}{1+\sqrt{\eta}},$$

donde $D = (\omega^{(2)}E + A_2)^2$. *Para el número de iteraciones* n *es válida la estimación*

$$n = n_0(\varepsilon) = \ln \varepsilon/(2 \ln \rho) \approx \ln \frac{1}{\varepsilon} \Big/ (4\sqrt{\eta}).$$

2. Problema de Dirichlet de diferencias para una ecuación elíptica de coeficientes variables. Examinemos un ejemplo de aplicación del método de las direcciones variables en el caso no conmutativo. Supongamos que sobre la red rectangular $\bar{\omega} = \{x_{ij} = (ih_1, jh_2) \in \bar{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha = l_\alpha/N_\alpha,\ \alpha = 1, 2,\}$ introducida en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$, se exige hallar la solución del siguiente problema de diferencias:

$$\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha(x)\, y_{\bar{x}_\alpha})_{x_\alpha} - q(x)\, y = -\varphi(x), \quad x \in \omega, \qquad (9)$$
$$y(x) = g(x), \quad x \in \gamma,$$

donde los coeficientes del esquema de diferencias satisfacen las condiciones

$$0 < c_1 \leqslant A_\alpha(x) \leqslant c_2, \quad 0 \leqslant d_1 \leqslant q(x) \leqslant d_2. \qquad (10)$$

En el espacio H de las funciones reticulares prefijadas sobre ω, con el producto escalar $(u, v) = \sum_{x \in \omega} u(x)\, v(x) h_1 h_2$, definiremos los operadores A_1 y A_2 de la siguiente forma:

$$A_\alpha y = -\Lambda_\alpha \overset{\circ}{y} = -(a_\alpha \overset{\circ}{y}_{x_\alpha})_{x_\alpha} + 0{,}5 q \overset{\circ}{y}, \qquad \alpha = 1, 2, \qquad y \in H,$$

$\overset{\circ}{y} \in \overset{\circ}{H}$, donde, como es usual, $\overset{\circ}{y}(x) = 0$ sobre γ.

Los operadores introducidos son autoconjugados en H y si $a_\alpha(x)$ depende únicamente de la variable x_α y $q(x)$ es constante, entonces los operadores A_1 y A_2 serán conmutables. En el caso general la conmutatividad no tendrá lugar, y para la resolución de la ecuación (1) correspondiente al problema de diferencias (9), se puede aplicar el método de las direcciones variables (3) examinado en el punto 1 del § 3.

El algoritmo del método tiene la sencilla forma

$$\omega^{(1)} y_{h+1/2} - \Lambda_1 y_{h+1/2} = \omega^{(1)} y_h + \Lambda_2 y_h + \varphi,$$
$$h_1 \leqslant x_1 \leqslant l_1 - h_1,$$
$$y_{h+1/2}(x) = g(x), \quad x_1 = 0,\ l_1,\ h_2 \leqslant x_2 \leqslant l_2 - h_2,$$
$$\omega^{(2)} y_{h+1} - \Lambda_2 y_{h+1} = \omega^{(2)} y_{h+1/2} + \Lambda_1 y_{h+1/2} + \varphi,$$
$$h_2 \leqslant x_2 \leqslant l_2 - h_2,$$
$$y_{h+1}(x) = g(x), \quad x_2 = 0,\ l_2, \quad h_1 \leqslant x_1 \leqslant l_1 - h_1.$$

Quedan por hallar las cotas δ_α y Δ_α de los operadores A_α, $\alpha = 1, 2$. Ya que las condiciones (10) se cumplen, entonces del lema 14 del § 2 del cap. V obtenemos

$$(y^2, 1)_{\omega_\alpha} \leqslant \varkappa_\alpha(x_\beta)(A_\alpha y, y)_{\omega_\alpha}, \quad \beta = 3 - \alpha, \quad \alpha = 1, 2,$$

donde $\varkappa_\alpha(x)_\beta = \max\limits_{x_\alpha \in \omega_\alpha} v^\alpha(x)$, y $v^\alpha(x)$ es la solución del siguiente problema de contorno tripuntual:

$$(A_\alpha v^\alpha_{\overline{x}_\alpha})_{x_\alpha} - 0{,}5qv^\alpha = -1, \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,$$
$$v^\alpha(x) = 0, \quad x_\alpha = 0,\ l_\alpha, \quad h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta.$$

El producto escalar según ω_α se define de la siguiente forma:

$$(u, v)_{\omega_\alpha} = \sum_{x_\alpha = h_\alpha}^{l_\alpha - h_\alpha} u(x) v(x) h_\alpha = \sum_{x_\alpha \in \omega_\alpha} u(x) v(x) h_\alpha.$$

Multiplicando (11) por h_β y sumando según x_β, obtenemos

$$\left(\frac{1}{\varkappa_\alpha} y^2, 1\right) \leqslant (A_\alpha y, y), \quad \alpha = 1, 2.$$

Por consiguiente, en calidad de δ_α se puede tomar

$$\delta_\alpha = \min_{h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta} \frac{1}{\varkappa_\alpha(x_\beta)}, \quad \beta = 3 - \alpha, \ \alpha = 1, 2.$$

Hallemos ahora Δ_α. Procederemos en analogía con el punto 2 del § 2. Designemos mediante $\mathscr{D}$ la parte diagonal de la matriz $\mathscr{A}_\alpha$, correspondiente al operador A_α:

$$\mathscr{D} y = d_\alpha(x) y,$$

$$d_\alpha(x) = \begin{cases} 0{,}5q(x) + \dfrac{1}{h_\alpha^2}(a_\alpha(x_\alpha, x_\beta) + a_\alpha(x_\alpha + h_\alpha, x_\beta)), \\ h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta. \end{cases}$$

Entonces tiene lugar la desigualdad

$$(A_\alpha y,\ y) \leqslant (2 - \lambda_{\min})(\mathscr{D} y,\ y) \leqslant (2 - \lambda_{\min}) \max_{x \in \omega} \times \\ \times\ d_\alpha\ (x)\ (y,\ y),$$

donde $\lambda_{\min}$ es la constante de la desigualdad operacional $\lambda_{\min} \mathscr{D} \leqslant A_\alpha$.

Hallemos $\lambda_{\min}$. Del lema 14 del § 2 del cap. V obtenemos

$$(d_\alpha y^2,\ 1)_{\omega_\alpha} \leqslant \rho_\alpha\ (x_\beta)\ (A_\alpha y,\ y)_{\omega_\alpha}, \qquad (12)$$

donde $\rho_\alpha\ (x_\beta) = \max\limits_{x_\alpha \in \omega_\alpha} w^\alpha (x)$, y $w^\alpha (x)$ es la solución del siguiente problema de contorno tripuntual:

$$(a_\alpha w^\alpha_{\overline{x}_\alpha})_{x_\alpha} - 0{,}5 q w^\alpha = -d_\alpha\ (x), \quad h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha,$$

$$w^\alpha\ (x) = 0, \quad x_\alpha = 0,\ l_\alpha, \quad h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta.$$

Multiplicando (12) por h_β y sumando según ω_β, obtendremos

$$\left(\frac{d_\alpha}{\rho_\alpha}\ y^2,\ 1\right) \leqslant (A_\alpha y,\ y), \quad \alpha = 1,\ 2.$$

Por consiguiente, en calidad de $\lambda_{\min}$ se puede tomar

$$\lambda_{\min} = \min_{h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta} \frac{1}{\rho_\alpha\ (x_\beta)},$$

por eso Δ_α es

$$\Delta_\alpha = \left(2 - \frac{1}{\max\limits_{x_\beta} \rho_\alpha\ (x_\beta)}\right) \max_{x \in \omega} d_\alpha\ (x), \quad \alpha = 1,\ 2.$$

Así pues, está hallada la información a priori necesaria para aplicar el método de las direcciones variables. Utilizando la condición (10), se puede mostrar, que la magnitud η, la cual define la velocidad de convergencia del método, para el ejemplo a examinar es $O\ (|\ h\ |^2)$, donde $|\ h\ |^2 = h_1^2 + h_2^2$. Por eso en virtud del teorema 2 para el número de iteraciones será válida la estimación

$$n = O\left(\frac{1}{|\ h\ |} \ln \frac{l}{\varepsilon}\right).$$

Examinemos un problema modelo. Sea el esquema de diferencias (9) prefijado sobre una red cuadrada en el cuadra-

do unidad ($N_1 = N_2 - N$, $l_1 = l_2 = 1$). Eligiremos los coeficientes $a_1(x)$, $a_2(x)$ y $q(x)$ de la siguiente forma:

$$a_1(x) = 1 + c\,[(x_1 - 0{,}5)^2 + (x_2 - 0{,}5)^2],$$

$$a_2(x) = 1 + c\,[0{,}5 - (x_1 - 0{,}5)^2 - (x_2 - 0{,}5)^2],$$

$$q(x) \equiv 0, \quad c > 0.$$

En este caso en las desigualdades (10) $c_1 = 1$, $c_2 = 1 + 0{,}5c$, $d_1 = d_2 = 0$, cambiando el parámetro c, obtendremos los coeficientes del esquema de diferencias (9) con diferentes características extremales.

Citemos el número de iteraciones para el método examinado de las direcciones variables en función de la relación c_2/c_1 y del número N de nodos por una dirección para $\varepsilon = 10^{-4}$.

Tabla 12

c_2/c_1	$N=32$	$N=64$	$N=128$
2	65	132	264
8	90	187	380
32	110	233	482
128	122	264	556
512	128	282	603

Comparemos este método con el método de relajación superior (véase el § 2 del cap. IX), con el método alternativo triangular (véase el § 2 del cap. X) y con el método implícito de Chebishev (véase el punto 3 del § 2 cap. VI). Por el número de iteraciones el método de las direcciones variables examinado cede ante el método de relajación superior y ante el método alternativo triangular, pero supera el método de Chebishev implícito en 1,5—2 veces. Sin embargo, por el volumen de trabajo computacional el método de las direcciones variables cederá también ante el método de Chebishev implícito.

Capítulo XII

Métodos de resolución de ecuaciones con operadores sin signo definido y degenerados

En el capítulo se estudian los métodos directos e iterativos de resolución de ecuaciones con un operador no degenerado y sin signo definido, con un operador complejo y también con un operador degenerado. En el § 1 para una ecuación con operador sin signo definido se examinan el método con los parámetros de Chebishev y el método de tipo variacional. En el § 2 para una ecuación con operador complejo de tipo especial están construidos los métodos de iteración simple y de las direcciones variables con parámetros iterativos complejos. En el § 3 son estudiados los métodos iterativos generales de resolución de ecuaciones con un operador degenerado, cuando el operador es no degenerado sobre la capa superior. El párrafo 4 está dedicado a la construcción de métodos directos e iterativos especiales para ecuaciones con un operador degenerado.

§ 1. Ecuaciones con un operador real sin signo definido

1. Esquema iterativo. Problema de la elección de los parámetros iterativos. Sea en el espacio de Hilbert H dada la ecuación

$$Au = f \tag{1}$$

con el operador lineal no degenerado A. Para resolver la ecuación (1) examinaremos el esquema iterativo implícito de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \tag{2}$$

con el operador no degenerado B y con $y_0 \in H$ arbitrario.

Los esquemas iterativos del tipo (2) se estudiaron en los capítulos VI y VIII, donde fueron propuestos algunos pro-

cedimientos para elegir los parámetros iterativos τ_k en función de las propiedades de los operadores A, B y D. Recordemos que D es un operador autoconjugado y definido positivo, el cual genera el espacio energético H_D. Fue mostrado que para la convergencia en H_D de los métodos iterativos examinados se exige la definición positiva del operador

$$C = D^{-1/2}(DB^{-1}A)D^{-1/2}. \tag{3}$$

Para operadores D concretos esta exigencia conduce a las siguientes condiciones para los operadores A y B:

1) el operador A debe ser definido positivo en H, si $D = A$, B o A^*B^{-1}A;

2) el operador B^*A debe ser definido positivo en H, si $D = A^*A$ o B^*B.

Existen problemas para los cuales estas exigencias no se cumplen, es decir, o bien el operador A no tiene signo definido o es difícil hallar un operador B, tal que B^*A sea un operador definido positivo. En calidad de ejemplo de tales problemas se puede citar el problema de Dirichlet para la ecuación de Helmgoltz en un rectángulo:

$$y_{\bar{x}_1x_1} + y_{\bar{x}_2x_2} + m^2y = 0, \quad x \in \omega,$$
$$y(x) = g(x), \quad x \in \gamma,$$

donde $m^2 > 0$.

Este párrafo está dedicado a la construcción de métodos iterativos de dos capas implícitas para el caso, cuando el operador C es un operador no degenerado y sin signo definido en H. Aquí nosotros examinaremos sólo operadores C reales, el caso complejo se estudia en el § 2.

Pasemos a la construcción de los métodos iterativos. En la ecuación

$$z_{k+1} = (E - \tau_{k+1}B^{-1}A)z_k, \quad k = 0, 1, \ldots, 2n-1,$$

para el error $z_k = y_k - u$ del esquema iterativo (2) haremos la sustitución $z_k = D^{-1/2}x_k$ y pasaremos a la ecuación para el error equivalente x_k:

$$x_{k+1} = (E - \tau_{k+1}C)x_k, \quad k = 0, 1, \ldots, 2n-1, \tag{4}$$

donde el operador C está definido en (3). Puesto que el operador C no tiene signo definido, entonces es obvio, que la norma del operador $E - \tau_{k+1}C$ será mayor o igual a la unidad para cualquier τ_{k+1}.

Examinemos ahora la ecuación que relaciona los errores en las iteraciones pares. De (4) obtenemos

$$x_{2k+2} = (E - \tau_{2k+2}C)(E - \tau_{2k+1}C)x_{2k}, \quad k = 0, 1, \ldots, \ldots, n-1. \quad (5)$$

Si se designa

$$\omega_{k+1} = -\tau_{2k+2}\tau_{2k+1}, \quad k = 0, 1, \ldots, n-1, \quad (6)$$

y se exige que los parámetros iterativos τ_{2k+2} y τ_{2k+1} para cualquier k satisfagan la relación

$$1/\tau_{2k+2} + 1/\tau_{2k+1} = 2\alpha, \quad k = 0, 1, \ldots, n-1, \quad (7)$$

donde α es una constante indefinida por ahora, entonces (5) se puede escribir en la forma

$$x_{2k+2} = (E - \omega_{k+1}\bar{C})\, x_{2k}, \quad k = 0, 1, \ldots, \bar{C} = C^2 - 2\alpha C. \quad (8)$$

Si ω_{k+1} y α son hallados, entonces los parámetros τ_{2k+2} y τ_{2k+1} en virtud de (6) y (7) se determinan por las fórmulas

$$\begin{aligned} \tau_{2k+1} &= -\alpha\omega_{k+1} - \sqrt{\alpha^2\omega_{k+1}^2 + \omega_{k+1}}, \\ \tau_{2k+2} &= -\alpha\omega_{k+1} + \sqrt{\alpha^2\omega_{k+1}^2 + \omega_{k+1}}, \\ k &= 0, 1, \ldots, n-1. \end{aligned} \quad (9)$$

De (8) obtenemos

$$\begin{aligned} x_{2n} &= \prod_{j=1}^{n} (E - \omega_j\bar{C})\, x_0, \\ \| x_{2n} \| &\leqslant \| \prod_{j=1}^{n} (E - \omega_j\bar{C}) \| \, \| x_0 \|. \end{aligned} \quad (10)$$

Ya que el operador $\bar{C}$ depende de α entonces el requisito de definición positiva del operador $\bar{C}$ será una de las condiciones, a la cual está subordinada la elección del parámetro α. Además, de (10) se desprende que los parámetros ω_j, $1 \leqslant j \leqslant n$, y el parámetro α se deben elegir de la condición de mínimo de la norma del operador de resolución $\prod_{j=1}^{n} (E - \omega_j\bar{C})$.

Este problema sobre la mejor elección de los parámetros iterativos ω_j y α, y por consiguiente, de los parámetros τ_k para el esquema (2) será resuelto más abajo. Primeramente

estableceremos la relación del procedimiento propuesto de construcción del método iterativo con el procedimiento, basado en la transformación de Gauss para el caso de un operador autoconjugado C.

Notemos, que el cambio $u = D^{-1/2}{}_{(x)}$, $f = BD^{-1/2}\varphi$ permite escribir la ecuación inicial (1) en la siguiente forma:

$$Cx = \varphi, \tag{11}$$

donde el operador C está definido en (3). Utilizando (11), obtendremos

$$\overline{C}x = C^2x - 2\alpha Cx = (C - 2\alpha E)\,\varphi = \overline{\varphi}. \tag{12}$$

Así sucesivamente, si designamos $v_k = \mathscr{D}^{1/2} y_k$, donde y_k es una aproximación iterativa en el esquema (2), entonces hallaremos fácilmente

$$x_k = D^{1/2}z_k = D^{1/2}y_k - D^{1/2}u = v_k - x.$$

Sustituyendo x_k en (8) y tomando en consideración (12), obtendremos el esquema iterativo

$$\frac{v_{2k+2} - v_{2k}}{\omega_{k+1}} + \overline{C}v_{2k} = \overline{\varphi}, \quad k = 0,\ 1,\ \ldots \tag{13}$$

De esta forma, el esquema (13) es un esquema de dos capas explícito para la ecuación transformada (12).

Sea $C = C^*$. Recordemos que en este caso la primera transformación de Gauss consiste en el paso de la ecuación (11) a la ecuación $\overline{\overline{C}}x = C^2x = C\varphi = \overline{\overline{\varphi}}$. Ya que C es un operador no degenerado, entonces el operador C^2 será definido positivo en H. Por eso la transformación indicada nos conduce a un operador con signo definido. Para resolver una ecuación con tal operador se puede utilizar un esquema de dos capas del tipo (13), sustituyendo $\overline{C}$ por $\overline{\overline{C}}$ y $\overline{\varphi}$ por $\overline{\overline{\varphi}}$. Es obvio, que tal método es un caso particular (para $\alpha = 0$) del método que se examina.

2. Transformación del operador en el caso autoconjugado. Supondremos que el operador C es autoconjugado en H. Entonces el operador $\overline{C} = C^2 - \alpha C$ también es autoconjugado en H. Nuestro objetivo más cercano será eligir el parámetro α de manera tal, que el operador $\overline{C}$ sea definido positivo, y hallar las cotas $\gamma_1 = \gamma_1(\alpha)$ y $\gamma_2 = \gamma_2(\alpha)$ de este operador, es decir, las magnitudes de las desigualdades

$$\gamma_1 E \leqslant \overline{C} \leqslant \gamma_2 E, \quad \gamma_1 > 0. \tag{14}$$

Si el valor indicado para α existe, entonces en virtud de la estimación

$$\left\| \prod_{j=1}^{n} (E - \omega_j \overline{C}) \right\| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} \left| \prod_{j=1}^{n} (1 - \omega_j t) \right|$$

el problema de encontrar los parámetros ω_j, $j = 1, 2, \ldots, n$, se reduce a la construcción del polinomio $P_n(t)$ de grado n, normado por la condición $P_n(0) = 1$ y con la menor desviación del cero en el segmento $[\gamma_1, \gamma_2]$ del semieje positivo. Este problema fue estudiado por nosotros anteriormente en el capítulo VI durante la construcción del método de Chebishev. La solución tiene la forma

$$\omega_k = \frac{\omega_0}{1 + \rho_0 \mu_k}, \quad \mu_k \in \mathfrak{M}_n^* = \left\{ -\cos \frac{2i-1}{2n} \pi, \; i = 1, 2, \ldots, n \right\},$$

donde $k = 1, 2, \ldots, n$,

$$\omega_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1 - \xi}{1 + \xi}, \quad \rho_1 = \frac{1 - \sqrt{\xi}}{1 + \sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Con esto en virtud de (10) para el error x_{2n} será válida la estimación

$$\| x_{2n} \| \leqslant q_n \| x_0 \|, \quad q_n = 2\rho_1^n / (1 + \rho_1^{2n}).$$

De aquí se sigue que la elección del parámetro α debe estar subordinada a la condición de máximo de la relación γ_1/γ_2.

Hallemos el valor óptimo para el parámetro α. Supongamos que los valores propios μ del operador C están contenidos en los intervalos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$. Como el operador C no tiene signo definido y es no degenerado entonces

$$\overset{\circ}{\gamma}_1 \leqslant \overset{\circ}{\gamma}_2 < 0 < \overset{\circ}{\gamma}_3 \leqslant \overset{\circ}{\gamma}_4. \tag{15}$$

Hallemos los valores propios λ del operador $\overline{C} = C^2 - 2\alpha C$. Es fácil ver, que los valores propios de los operadores $\overline{C}$ y C están enlazados por la relación:

$$\lambda = \mu^2 - 2\alpha\mu, \quad \mu \in \Omega, \tag{16}$$

donde Ω consiste de los dos intervalos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$.

Hallemos primeramente las restricciones sobre α que aseguran la positividad de los valores propios λ, es decir, la definición positiva del operador $\overline{C}$. Analizando la desigualdad $\mu^2 - 2\alpha\mu > 0$, encontramos que ella tiene lugar para μ que varía fuera del intervalo $[0, 2\alpha]$. Por eso esta desigual-

dad se cumplirá para $\mu \in \Omega$, si α satisface la condición

$$\overset{\circ}{\gamma}_2 \leqslant 2\alpha \leqslant \overset{\circ}{\gamma}_3. \qquad (17)$$

Consideraremos que (17) está cumplida. De (16) obtenemos que la transformación $\lambda = \lambda(\mu) = \mu^2 - 2\alpha\mu$ aplica el segmento $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ sobre el segmento $[\lambda_2, \lambda_1[$, y el segmento $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$, sobre el segmento $[\lambda_3, \lambda_4[$, donde $\lambda_i = \lambda(\overset{\circ}{\gamma}_i)$, $1 \leqslant i \leqslant 4$. De esta forma, todos los valores propios del operador C son positivos y están distribuidos sobre los segmentos $[\lambda_2, \lambda_1) \cup [\lambda_3, \lambda_4[$. Por eso en las desigualdades (14) conviene poner que

$$\gamma_1 = \text{mín}(\lambda_2, \lambda_3), \qquad \gamma_2 = \text{máx}(\lambda_1, \lambda_4).$$

Elijamos ahora $2\alpha \in (\overset{\circ}{\gamma}_2, \overset{\circ}{\gamma}_3)$ de la condición de máximo de la relación γ_1/γ_2. De (18) obtenemos

$$\gamma_1 = \begin{cases} \lambda_2 = \overset{\circ}{\gamma}_2(\overset{\circ}{\gamma}_2 - 2\alpha), & \overset{\circ}{\gamma}_2 < 2\alpha \leqslant \overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3, \\ \lambda_3 = \overset{\circ}{\gamma}_3(\overset{\circ}{\gamma}_3 - 2\alpha), & \overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3 \leqslant 2\alpha < \overset{\circ}{\gamma}_3, \end{cases}$$

$$\gamma_2 = \begin{cases} \lambda_4 = \overset{\circ}{\gamma}_4(\overset{\circ}{\gamma}_4 - 2\alpha), & 2\alpha \leqslant \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4, \\ \lambda_1 = \overset{\circ}{\gamma}_1(\overset{\circ}{\gamma}_1 - 2\alpha), & \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4 \leqslant 2\alpha. \end{cases}$$

Introduzcamos las siguientes notaciones:

$$\Delta_1 = \overset{\circ}{\gamma}_2 - \overset{\circ}{\gamma}_1, \quad \Delta_2 = \overset{\circ}{\gamma}_4 - \overset{\circ}{\gamma}_3$$

y examinemos dos casos.

1) Sea primeramente $\Delta_1 \leqslant \Delta_2$, es decir, $\overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3 \leqslant \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4$. En este caso para $\varepsilon = \gamma_1/\gamma_2$ obtenemos la siguiente expresión:

$$\xi = \xi(\alpha) = \begin{cases} \dfrac{\overset{\circ}{\gamma}_2(\overset{\circ}{\gamma}_2 - 2\alpha)}{\overset{\circ}{\gamma}_4(\overset{\circ}{\gamma}_4 - 2\alpha)}, & \overset{\circ}{\gamma}_2 < 2\alpha \leqslant \overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3, \\ & \text{crece respecto de } \alpha, \\ \dfrac{\overset{\circ}{\gamma}_3(\overset{\circ}{\gamma}_3 - 2\alpha)}{\overset{\circ}{\gamma}_4(\overset{\circ}{\gamma}_4 - 2\alpha)}, & \overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3 \leqslant 2\alpha \leqslant \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4, \\ & \text{decrece respecto de } \alpha, \\ \dfrac{\overset{\circ}{\gamma}_3(\overset{\circ}{\gamma}_3 - 2\alpha)}{\overset{\circ}{\gamma}_1(\overset{\circ}{\gamma}_1 - 2\alpha)}, & \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4 \leqslant 2\alpha, \\ & \text{decrece respecto de } \alpha. \end{cases}$$

Por consiguiente, en este caso el valor óptimo de α es

$$\alpha = \alpha_0 = (\overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3)/2, \tag{19}$$

y a su vez está cumplida la condición (17). Para $\alpha = \alpha_0$ tenemos

$$\gamma_1 = \lambda_2 = \lambda_3 = -\overset{\circ}{\gamma}_2 \overset{\circ}{\gamma}_3, \tag{20}$$

$$\gamma_2 = \lambda_4 = \overset{\circ}{\gamma}_4 (\Delta_2 - \Delta_1) - \overset{\circ}{\gamma}_1 \overset{\circ}{\gamma}_4 \geqslant \lambda_1. \tag{21}$$

2) Sea ahora $\Delta_1 \geqslant \Delta_2$, es decir, $\overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3 \geqslant \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4$. En este caso tendremos

$$\xi = \xi(\alpha) = \begin{cases} \dfrac{\overset{\circ}{\gamma}_3 (\overset{\circ}{\gamma}_3 - 2\alpha)}{\overset{\circ}{\gamma}_4 (\overset{\circ}{\gamma}_4 - 2\alpha)}, & 2\alpha \leqslant \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4, \\ & \text{crece respecto a } \alpha \\ \dfrac{\overset{\circ}{\gamma}_2 (\overset{\circ}{\gamma}_2 - 2\alpha)}{\overset{\circ}{\gamma}_1 (\overset{\circ}{\gamma}_1 - 2\alpha)}, & \overset{\circ}{\gamma}_1 + \overset{\circ}{\gamma}_4 \leqslant 2\alpha \leqslant \overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3, \\ & \text{crece respecto a } \alpha, \\ \dfrac{\overset{\circ}{\gamma}_2 (\overset{\circ}{\gamma}_2 - 2\alpha)}{\overset{\circ}{\gamma}_1 (\overset{\circ}{\gamma}_1 - 2\alpha)}, & \overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3 \leqslant 2\alpha < \overset{\circ}{\gamma}_3, \\ & \text{decrece respecto a } \alpha. \end{cases}$$

Por lo tanto, también en este caso el valor óptimo del parámetro α se determina por la fórmula (19), el valor de γ_1 está dado en (20), mientras que

$$\gamma_2 = \lambda_1 = \overset{\circ}{\gamma}_1 (\Delta_2 - \Delta_1) - \overset{\circ}{\gamma}_1 \overset{\circ}{\gamma}_4 \geqslant \lambda_4. \tag{22}$$

Así, está demostrado

LEMA 1. *Sean contenidos los valores propios del operador C en los intervalos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$, $\overset{\circ}{\gamma}_2 < 0 < \gamma_3$. Entonces para el operador $\overline{C} = C^2 - \alpha C$ si $\alpha = \alpha_0 = (\overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3)/2$ son válidas las desigualdades*

$$\gamma_1 E \leqslant \overline{C} \leqslant \gamma_2 E, \quad \gamma_1 > 0,$$

donde

$$\gamma_1 = -\overset{\circ}{\gamma}_2 \overset{\circ}{\gamma}_3, \quad \gamma_2 = \max [\overset{\circ}{\gamma}_4 (\Delta_2 - \Delta_1),$$

$$\overset{\circ}{\gamma}_1 (\Delta_2 - \Delta_1)] - \overset{\circ}{\gamma}_1 \overset{\circ}{\gamma}_4$$

Para el valor de α indicado la razón γ_1/γ_2 es máxima.

Las afirmaciones del lema se deducen de (19)—(22). Señalemos que $\alpha_0 = 0$, únicamente en el caso, cuando $\overset{\circ}{\gamma}_2 = -\overset{\circ}{\gamma}_3$.

3. Método iterativo con parámetros de Chebishev. Más arriba hemos examinado el esquema iterativo de dos capas (2) para el cual los parámetros τ_k, $k = 1, 2, \ldots, 2n$ se expresan mediante $\bar{\omega}_k$, $1 \leqslant k \leqslant n$ y α por las fórmulas (9). Con esto los parámetros $\bar{\omega}_k$ son los parámetros iterativos del método de Chebishev y se determinan por las fórmulas respectivas, y la información apriorística necesaria para esto al igual que el valor óptimo del parámetro α están dados en lema 1.

Notemos que nosotros supusimos la pertenencia de los valores propios μ del operador autoconjugado C a los segmentos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$. De la definición (3) del operador C se sigue que los valores propios del operador C son también al mismo tiempo valores propios del problema siguiente:

$$Au - \mu Bu = 0. \tag{23}$$

Para cerciorarse de esto, es suficiente multiplicar el primer miembro de esta ecuación por el operador $D^{1/2}B^{-1}$ y hacer un cambio, poniendo $u = D^{-1/2}v$. Observemos que el operador C será autoconjugado en H, si es autoconjugado el operador $DB^{-1}A$. Formulemos los resultados obtenidos como un teorema.

TEOREMA 1. *Sea el operador $DB^{-1}A$ autoconjugado en H y supongamos que los valores propios del problema (23) pertenecen a los segmentos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$, $\overset{\circ}{\gamma}_1 \leqslant \overset{\circ}{\gamma}_2 < 0 < \overset{\circ}{\gamma}_3 \leqslant \overset{\circ}{\gamma}_4$. Para el proceso iterativo (2) con los parámetros*

$$\tau_{2k-1} = -\alpha_0\omega_k - \sqrt{\alpha_0^2\omega_k^2 + \omega_k}, \quad \tau_{2k} = -\alpha_0\omega_k + \sqrt{\alpha_0^2\omega_k^2 + \omega_k},$$

$k = 1, 2, \ldots, n$,

es válida la estimación

$$\| z_{2n} \|_D \leqslant q_n \| z_0 \|_D,$$

donde

$$\omega_k = \frac{\omega_0}{1 + \rho_0\mu_k}, \quad \mu_k \in \mathfrak{M}_n^* =$$

$$= \left\{ -\cos\frac{(2i-1)\pi}{2n}, \ 1 \leqslant i \leqslant n \right\}, \quad 1 \leqslant k \leqslant n,$$

$$\omega_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}},$$

$$q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}} \quad \xi = \frac{\gamma_1}{\gamma_2},$$

$$\alpha_0 = 0{,}5\,(\overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3), \quad \gamma_1 = -\overset{\circ}{\gamma}_2\overset{\circ}{\gamma}_3,$$

$$\gamma_2 = \text{máx}\,[\overset{\circ}{\gamma}_4\,(\Delta_2 - \Delta_1), \quad \overset{\circ}{\gamma}_1\,(\Delta_2 - \Delta_1)] - \overset{\circ}{\gamma}_1\overset{\circ}{\gamma}_4,$$

$$\Delta_1 = \overset{\circ}{\gamma}_2 - \overset{\circ}{\gamma}_1, \quad \Delta_2 = \overset{\circ}{\gamma}_4 - \overset{\circ}{\gamma}_3.$$

El método iterativo (2) con los parámetros indicados τ_k lo llamaremos *método de Chebishev.*

Examinemos algunos casos particulares. Sea $\Delta_1 = \Delta_2$, es decir, las longitudes de los segmentos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$ son iguales. En este caso tenemos

$$\gamma_1 = -\overset{\circ}{\gamma}_2\overset{\circ}{\gamma}_3, \quad \gamma_2 = -\overset{\circ}{\gamma}_1\overset{\circ}{\gamma}_4, \quad \xi = \frac{\overset{\circ}{\gamma}_2\overset{\circ}{\gamma}_3}{\overset{\circ}{\gamma}_1\overset{\circ}{\gamma}_4}.$$

Mostremos que en el caso que se examina la lista construida de los parámetros τ_k es la mejor. Esta afirmación es necesario demostrarla, por cuanto durante la construcción de los parámetros τ_k para el esquema (2) nosotros impusimos n condiciones (7), y por consiguiente, la elección de los parámetros estuvo subordinada a restricciones complementarias.

De (5) y (8) hallaremos que

$$x_{2n} = Q_{2n}(C)x_0 = P_n(\overline{C})x_0,$$

donde

$$Q_{2h}(C) = \prod_{j=1}^{2n}(E - \tau_j C) = P_n(\overline{C}) = \prod_{j=1}^{n}(E - \omega_j \overline{C}). \qquad (24)$$

Examinemos los correspondientes polinomios algebraicos $Q_{2n}(\mu)$ y $P_n(\lambda)$ ($\lambda = \mu^2 - 2\alpha\mu$). Si los parámetros ω_j son elegidos de la manera indicada en el teorema 1, entonces el politnomio $P_n(\lambda)$ se expresa de modo siguiente mediante el polinomio de Chebishev de 1-er género $T_n(\lambda)$ (véase el punto 1 del § 2 del cap. VI):

$$P_n(\lambda) = q_n T_n\left(\frac{1-\omega_0\lambda}{\rho_0}\right), \quad P_n(0) = 1,$$

$$\max_{\gamma_1 \leqslant \lambda \leqslant \gamma_2} |P_n(\lambda)| = q_n.$$

Notemos que en los puntos $\gamma_1 = \lambda_0 < \lambda_1 < \ldots < \lambda_n = \gamma_2$, donde

$$\lambda_k = \frac{1 - \rho_0 \cos \frac{k\pi}{n}}{\omega_0}, \quad k = 0, 1, \ldots, n,$$

el polinomio $P_n(\lambda)$ alcanza los valores extremos sobre $[\gamma_1, \gamma_2]$:

$$P_n(\lambda_k) = (-1)^k q_n, \quad k = 0, 1, \ldots, n. \tag{25}$$

Puesto que en virtud de (24) tiene lugar la igualdad $Q_{2n}(\mu) = P_n(\lambda)$, donde λ y μ están enlazados por la relación $\lambda = \mu^2 - 2\alpha\mu$, entonces de (25) hallaremos

$$Q_{2n}(\mu_k^-) = Q_{2n}(\mu_k^+) = (-1)^k q_n, \quad k = 0, 1, \ldots, n, \tag{26}$$

donde μ_k^- y μ_k^+ son las raíces de la ecuación cuadrática

$$\mu_k^2 - 2\alpha\mu_k - \lambda_k = 0, \quad k = 0, 1, \ldots, n. \tag{27}$$

Luego, para el caso a examinar la transformación $\lambda = \lambda(\mu) = \mu^2 - 2\alpha\mu$ aplicada cada uno de los segmentos $[\gamma_1, \gamma_2]$ y $[\gamma_3, \gamma_4]$ sobre un segmento $[\gamma_1, \gamma_2]$. Al mismo tiempo a los puntos $\mu = \overset{\circ}{\gamma}_2$ y $\mu = \overset{\circ}{\gamma}_3$ les corresponde $\lambda = \gamma_1$, y a $\mu = \overset{\circ}{\gamma}_1$, y $\mu = \overset{\circ}{\gamma}_4$ les corresponde $\lambda = \gamma_2$. Por lo tanto las raíces de la ecuación (27) están dispuestas de la manera siguiente:

$$\overset{\circ}{\gamma}_1 = \mu_n^- < \mu_{n-1}^- < \ldots < \mu_0^- = \overset{\circ}{\gamma}_2,$$

$$\gamma_3 = \mu_0^+ < \mu_1^+ < \ldots < \mu_n^+ = \overset{\circ}{\gamma}_4.$$

Supongamos ahora que la lista de los parámetros τ_k construida en el teorema 1 no es la mejor. Esto significa que existe otro polinomio de grado no superior a $2n$ de la forma

$$\overline{Q}_{2n}(\mu) = \prod_{j=1}^{2n} (1 - \overline{\tau}_j \mu),$$

para el cual

$$\max_{\mu \in \Omega} |\overline{Q}_{2n}(\mu)| < q_n, \quad \Omega = [\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2] \cup [\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4].$$

Examinemos la diferencia $R_{2n}(\mu) = Q_{2n}(\mu) - \overline{Q}_{2n}(\mu)$, la cual es un polinomio de grado no superior a $2n$. La demostración de la existencia de $2n + 2$ raíces del polinomio $R_{2n}(\mu)$ nos conduce a la conclusión sobre la inexactitud de la suposición hecha más arriba.

Para la demostración examinemos los valores de $R_{2n}(\mu)$ en los puntos μ_k^-, $0 \leqslant k \leqslant n$. Ya que, según la suposición, $-q_n < \overline{Q}_{2n}(\mu) < q_n$, $\mu \in \Omega$, entonces

$$R_{2n}(\mu_k^-) = Q_{2n}(\mu_k^-) - \overline{Q}_{2n}(\mu_k^-) = (-1)^k q_n - \overline{Q}_{2n}(\mu_k^-)$$

y $R_{2n}(\mu_k^-) < 0$, si k es impar y $R_{2n}(\mu_k^-) > 0$, si k es par. Por lo tanto, al pasar de μ_k^- a μ_{k+1}^-, $0 \leqslant k \leqslant n-1$, el polinomio $R_{2n}(\mu)$ cambia de signo. Por eso en el segmento $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ existen n raíces de este polinomio. Análogamente, examinando los valores de $R_{2n}(\mu)$ en los puntos μ_k^+, $0 \leqslant k \leqslant n$, demostraremos la existencia de n raíces también sobre el segmento $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$. Después, puesto que

$$R_{2n}(\overset{\circ}{\gamma}_2) = R_{2n}(\mu_0^-) > 0, \quad R_{2n}(\overset{\circ}{\gamma}_3) = R_{2n}(\mu_0^+) > 0,$$
$$R_{2n}(0) = 0,$$

entonces en el intervalo $(\overset{\circ}{\gamma}_2, \overset{\circ}{\gamma}_3)$ hay o bien dos raíces diferentes (una de las cuales es el cero) del polinomio $R_{2n}(\mu)$, o el cero es una raíz múltiple. Por consiguiente, sobre el segmento $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_4]$ el polinomio $R_{2n}(\mu)$ posee $2n + 2$ raíces, lo cual no es posible.

Así pues, para el caso $\Delta_1 = \Delta_2$ la lista de los parámetros τ_k construida en el teorema 1 es la mejor.

Sea ahora $\Delta_1 \leqslant \Delta_2$. En este caso tenemos $\gamma_1 = -\overset{\circ}{\gamma}_2\overset{\circ}{\gamma}_3$ y $\gamma_2 = \overset{\circ}{\gamma}_4(\Delta_2 - \Delta_1) - \overset{\circ}{\gamma}_1\overset{\circ}{\gamma}_4$. Ya que $\gamma_2 = \overset{\circ}{\gamma}_4(\Delta_2 - \Delta_1) - \overset{\circ}{\gamma}_1\overset{\circ}{\gamma}_4 = \overset{\circ}{\gamma}_4(\overset{\circ}{\gamma}_4 - \overset{\circ}{\gamma}_3 - \overset{\circ}{\gamma}_2)$ entonces γ_1 y γ_2 no dependen de $\overset{\circ}{\gamma}_1$. Por lo tanto, para cualquier $\overset{\circ}{\gamma}_1$ del intervalo $\overset{\circ}{\gamma}_2 + \overset{\circ}{\gamma}_3 - \gamma_4 \leqslant \overset{\circ}{\gamma}_1 \leqslant \overset{\circ}{\gamma}_2$ tenemos una misma lista de parámetros τ_k, y el método interativo (2) converge con igual velocidad para todo $\overset{\circ}{\gamma}_1$ del intervalo indicado.

Como conclusión notemos que el conjunto de los parámetros τ_k construido en el teorema 1 será el mejor posible también para el caso, cuando $n = 1$, y Δ_1 y Δ_2 no son obligatoriamente iguales. Este es el caso del método cíclico de iteración

simple, para el cual en el esquema (2) $\tau_{2k-1} \equiv \tau_1$, $\tau_{2k} = \tau_2$, $k = 1, 2, \ldots$, y τ_1, τ_2 se encuentran por las fórmulas del teorema 1 para $n = 1$ ($\omega_1 = \omega_0$)

$$\tau_1 = -\alpha_0\omega_0 - \sqrt{\alpha_0^2\omega_0^2 + \omega_0}, \quad \tau_2 = -\alpha_0\omega_0 + \sqrt{\alpha_0^2\omega_0^2 + \omega_0},$$

donde $\omega_0 = 2/(\gamma_1 + \gamma_2)$. Ya que en este caso tenemos

$$x_{2n} = \prod_{j=1}^{n} (E - \omega_0\overline{C})\, x_0 = (E - \omega_0\overline{C})^n x_0,$$

$$\| E - \omega_0\overline{C} \| \leqslant \rho_0, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

entonces para el error z_{2n} del esquema iterativo (2) tendremos la estimación

$$\| z_{2n} \|_D \leqslant \rho_0^n \| z_0 \|_D.$$

Ya que en virtud de (6) y (7) dos parámetros τ_1 y τ_2 se cambian por los parámetros ω_1 y α, y los últimos se eligen de manera óptima ($\omega_1 = \omega_0$, $\alpha = \alpha_0$), entonces realmente, los parámetros τ_1 y τ_2 para el método de iteración simple están elegidos mejores.

4. Métodos iterativos de tipo variacional. Más arriba hemos examinado los métodos iterativos para el caso del operador autoconjugado $DB^{-1}A$, cuando no todos los valores propios del problema (23) son del mismo signo. Con esto la convergencia del método iterativo (2) fue garantizada por la construcción de un conjunto especial de parámetros iterativos. Examinemos ahora los métodos iterativos del tipo (2), cuya convergencia para la elección usual de los parámetros iterativos está asegurada por la estructura del operador B. Con este procedimiento de construcción de los esquemas iterativos nosotros tuvimos relación en el método de simetrización de una ecuación (véase el cap. VI, § 4, punto 4) y durante el estudio de los métodos de los errores mínimos y de los errores conjugados en el capítulo VIII.

Supongamos que el operador B tiene la forma

$$B = (A^*)^{-1}\widetilde{B}, \tag{28}$$

donde $\widetilde{B}$ es un operador autoconjugado y definido positivo arbitrario. En calidad del operador D tomaremos $\widetilde{B}$. Entonces $DB^{-1}A = A^*A$, $C = \widetilde{B}^{-1/2}A^*A\widetilde{B}^{-1/2}$. Si el operador B no es degenerado y no tiene signo definido, entonces el ope-

rador C es igualmente definido positivo. Además, el operador C es autoconjugado en H. Por eso, si son prefijadas γ_1 y γ_2 en las desigualdades $\gamma_1 E \leqslant C \leqslant \gamma_2 E$, $\gamma_1 > 0$ o en las desigualdades a ellas equivalentes

$$\gamma_1 \widetilde{B} < A^* A \leqslant \gamma_2 \widetilde{B}, \quad \gamma_1 > 0, \tag{29}$$

entonces los parámetros τ_k en (2) pueden elegirse por medio de las fórmulas del método de Chebishev de dos capas (véase el cap. VI, § 2, punto 1)

$$\tau_k = \frac{\tau_0}{1+\rho_0 \mu_k}, \quad \mu_k \in \mathfrak{M}_n^* = \left\{ -\cos \frac{(2i-1)\pi}{2n}, \ i \leqslant i \leqslant n \right\},$$
$$k = 1, 2, \ldots, n, \tag{30}$$
$$\tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

De esta forma, tiene lugar

TEOREMA 2. *Sea A un operador no degenerado. Para el método iterativo (2), (28) con los parámetros (30), donde γ_1 y γ_2 están dados prefijados en (29), tiene lugar la estimación*

$$\| z_n \|_{\widetilde{B}} \leqslant q_n \| z_0 \|_{\widetilde{B}}, \quad q_n = \frac{2\rho_1^n}{1+\rho_1^{2n}}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}.$$

Si las constantes γ_1 y γ_2 en (29) o bien son desconocidas, o pueden ser estimadas demasiado burdamente, entonces se pueden aprovechar los métodos iterativos de tipo variacional examinados en el capítulo VIII.

Si para el esquema (2), (28) se eligen los parámetros τ_k por las fórmulas

$$\tau_{k+1} = \frac{(r_k, r_k)}{(Aw_k, r_k)}, \quad k = 0, 1, \ldots,$$

donde $r_k = Ay_k - f$ es el defecto y w_k, la corrección definida de la ecuación $\widetilde{B} w_k = A^* r_k$, entonces obtendremos el método de los errores mínimos (véase el cap. VIII, § 2, punto 4). Como es conocido, para el error z_n de este método tiene lugar la estimación

$$\| z_n \|_{\widetilde{B}} \leqslant \rho_0^n \| z_0 \|_{\widetilde{B}},$$

donde ρ_0 está definido en (30).

Si examinamos el esquema iterativo de tres capas

$$By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)y_k + (1 - \alpha_{k+1})By_{k-1} + \\ + \tau_{k+1}\alpha_{k+1}f, \quad k \geqslant 1,$$

$$By_1 = (B - \tau_1 A)y_0 + \tau_1 f, \quad y_0 \in H,$$

donde el operador B está definido en (28), y elegimos los parámetros iterativos α_{k+1} y τ_{k+1} mediante las fórmulas

$$\tau_{k+1} = \frac{(r_k, r_k)}{(Aw_k, r_k)}, \quad k = 0, 1, \ldots,$$

$$\alpha_{k+1} = \left(1 - \frac{\tau_{k+1}}{\tau_k}\frac{(r_k, r_k)}{(r_{k-1}, r_{k-1})}\frac{1}{\alpha_k}\right)^{-1}, \quad k = 1, 2, \ldots, \quad \alpha_1 = 1,$$

entonces obtendremos el método de los errores conjugados (véase el cap. VIII, § 4, punto 1). Para el error de este método es cierta la estimación

$$\| z_n \|_{\widetilde{B}} \leqslant q_n \| z_0 \|_{\widetilde{B}}.$$

5. Ejemplos. Examinenos la aplicaciön de los métodos construidos más arriba para encontrar la solución del problema de Dirichlet de diferencias para la ecuación de Helmgoltz en un rectángulo:

$$y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + m^2 y = -f(x), \quad x \in \omega,$$
$$y(x) = g(x), \quad x \in \gamma, \tag{31}$$

donde $\overline{\omega} = \{x_{ij} = (ih_1, jh_2),\ 0 \leqslant i \leqslant N_1,\ 0 < j \leqslant N_2,$ $h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2\}$, y γ es la frontera de la red $\overline{\omega}$.

Reduciremos el problema (31) a la ecuación operacional (1). En este caso H es el espacio de las funciones reticulares, definidas sobre ω con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x)\, v(x)\, h_1 h_2, \quad u \in H, \ v \in H.$$

Definamos el operador R de la siguiente forma: $Ry = -\Lambda \mathring{y}$, $y \in H$, $\mathring{y} \in \mathring{H}$ y $y(x) = \mathring{y}(x)$, $x \in \omega$, donde $\mathring{H}$ es el conjunto de las funciones reticulares, definidas sobre $\overline{\omega}$ y que se anulan sobre γ, y Λ es el operador de Laplace de diferencias $\Lambda y = y_{x_1 x_1} + y_{x_2 x_2}$. Entonces el operador A se

define por la igualdad $A = R - m^2E$. Ya que el operador R es autoconjugado en H y tiene valores propios

$$\lambda_k = \lambda_{k_1}^{(1)} + \lambda_{k_2}^{(2)}, \quad \lambda_{k_\alpha}^{\alpha} = \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{k_\alpha \pi h_\alpha}{2l_\alpha},$$

$$l \leqslant k_\alpha \leqslant N_\alpha - 1,$$

entonces el operador A también es autoconjugado en H y sus valores propios μ_k se expresan mediante λ_k por medio de la fórmula

$$\mu_k = \lambda_k - m^2, \quad k = (k_1, k_2), \quad 1 \leqslant k_\alpha \leqslant N_\alpha - 1, \quad \alpha = 1, 2. \tag{32}$$

Supongamos que m^2 no coincide con ningún λ_k. Designamos mediante λ_{m_1} y λ_{m_2} los valores propios λ_k más cercanos a m^2 respectivamente por debajo y por encima, es decir,

$$\lambda m_1 < m2 < \lambda_{m_2}. \tag{33}$$

En este caso el operador A no es degenerado y no tiene signo definido.

Para resolver la ecuación (1) con el operador A indicado examinaremos el esquema iterativo explícito (2) ($B = E$). Si suponemos $D = E$, entonces el operador $DB^{-1}A$ coincide con A y es autoconjugado en H. La elección de los parámetros iterativos en este caso se puede realizar utilizando el teorema 1. De (23) obtenemos que la información apriorística necesaria se da mediante las fronteras de los segmentos $[\overset{\circ}{\gamma}_1, \overset{\circ}{\gamma}_2]$ y $[\overset{\circ}{\gamma}_3, \overset{\circ}{\gamma}_4]$ de los semiejes positivo y negativo, sobre los cuales están distribuidos los valores propios del operador A.

De (32) y (33) hallaremos $\overset{\circ}{\gamma}_1 = \delta - m^2$, $\overset{\circ}{\gamma}_2 = \lambda_{m_1} - m^2$, $\overset{\circ}{\gamma}_3 = \lambda_{m_2} - m_2$, $\overset{\circ}{\gamma}_4 = \Delta - m^2$, donde

$$\delta = \min_k \lambda_k = \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha},$$

$$\Delta = \max_k \lambda_k = \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2l_\alpha}.$$

Hallemos ahora γ_1, γ_2 y la magnitud $\sqrt{\xi}$ la cual determina el número de iteraciones para el método a examinar, de for-

ma tal que $n \geqslant n_0(\varepsilon) = \ln(2/\varepsilon)/(2\sqrt{\xi})$. De las fórmulas del teorema 1 hallaremos

$$\gamma_1 = (m^2 - \lambda_{m_1})(\lambda_{m_2} - m^2),$$

$$\gamma_2 = \begin{cases} (\Delta - m^2)(\Delta + m^2 - \lambda_{m_1} - \lambda_{m_2}), & \lambda_{m_1} + \lambda_{m_2} \leqslant (\Delta + \delta), \\ (m^2 - \delta)(\lambda_{m_1} + \lambda_{m_2} - m^2 - \delta), & \lambda_{m_1} + \lambda_{m_2} \geqslant (\Delta + \delta). \end{cases}$$

La relación $\xi = \gamma_1/\gamma_2$ depende de m^2. Para obtener una representación sobre la calidad del método iterativo que se examina, hallaremos el valor m^2 del intervalo $(\lambda_{m_1}, \lambda_{m_2})$ para el cual ξ es máximo. Obtendremos

$$m^2 = 0{,}5(\lambda_{m_1} + \lambda_{m_2}),$$

y con esto

$$\gamma_1 = \left(\frac{\lambda_{m_2} - \lambda_{m_1}}{2}\right)^2,$$

$$\gamma_2 = \begin{cases} (\Delta - m^2)^2, & 2m^2 \leqslant \Delta + \delta, \\ (m^2 - \delta)^2, & 2m^2 \geqslant \Delta + \delta. \end{cases}$$

Si m^2 es pequeño, o sea λ_{m_1} y λ_{m_2} son cercanos a δ, entonces $\gamma_1 = O(1)$, y $\gamma_2 = (\Delta - m^2)^2 = O\left(\frac{1}{|h|^4}\right)$. En este caso el mejor $\xi = O(|h|^4)$. Si λ_{m_1} y λ_{m_2} son cercanos a Δ, entonces de nuevo obtendremos $\xi = O(|h|^4)$. Solamente para el caso, cuando λ_{m_1} y λ_{m_2} son próximos a $0{,}5(\Delta + \delta)$, obtendremos

$$\gamma_1 = O\left(\frac{1}{|h|^2}\right) \text{ y } \gamma_2 = O\left(\frac{1}{|h|^2}\right),$$

así que

$$\xi = O(|h|^2).$$

Observemos que el problema de diferencias (31) puede ser resuelto por uno de los métodos directos, examinados por nosotros en los capítulos III y IV: o bien por el método de reducción completa, o por el método de separación de variables. Los problemas de contorno tripuntuales que aparecen con éste deben resolverse, a diferencia del caso $m = 0$, por el método de factorización no monótona.

§ 2. Ecuaciones con un operador complejo

1. Método de iteración simple. Sea en el espacio de Hilbert H dada la ecuación

$$Au + qu = f, \tag{1}$$

donde A es un operador hermitiano, y $q = q_1 + iq_2$ es un número complejo. Para la resolución aproximada de la ecuación (1) examinaremos el esquema explícito de dos capas

$$\frac{y_{k+1} - y_k}{\tau} + (A + qE)\, y_k = f, \quad k = 0, 1, \ldots, \; y_0 \in H, \tag{2}$$

donde $\tau = \tau_1 + i\tau_2$ es un parámetro iterativo complejo.

Supondremos que $q_1 \neq 0$, y γ_1, γ_2 son los constantes en las desigualdades

$$\gamma_1 E \leqslant A \leqslant \gamma_2 E. \tag{3}$$

Investiguemos la convergencia del esquema iterativo (2) en el espacio energético H ($D = E$) y hallemos el valor óptimo para el parámetro iterativo τ. Utilizando (1) y (2), escribamos la ecuación para el error $x_k = y_k - u$ en la forma:

$$x_{k+1} = s x_k, \quad k = 0, 1, \ldots, \; s = E - \tau C, \tag{4}$$

donde

$$C = A + qE.$$

De (4) hallaremos

$$x_n = S^n x_0, \quad \| x_n \| \leqslant \| S^n \| \, \| x_0 \|. \tag{5}$$

Estudiemos el operador de transición de una iteración a otra. Puesto que el operador A es hermitiano, entonces

$$C^* = A + \bar{q}E, \quad C^*C = CC^*,$$

es decir, el operador C es un operador normal. Por eso también es normal el operador S. Es conocido (véase el cap. V, § 1 punto 2), que para un operador normal S son válidas las siguientes relaciones.

$$\| S^n \| = \| S \|^n, \quad \| S \| = \sup_{x \neq 0} \frac{|(Sx, x)|}{(x, x)}.$$

Por lo tanto, de (5) se deduce que el problema de la elección del parámetro iterativo τ se reduce a encontrar el mismo de la condición de mínimo de la norma del operador S.

Resolvamos este problema. De (3) se deducirá que

$$z = \frac{(Cx, x)}{(x, x)} \in \Omega,$$

$$\Omega = \{z = z_1 + a(z_2 - z_1), \quad 0 \leqslant a \leqslant 1, \quad z_1 = \gamma_1 + q, \quad z_2 = \gamma_2 + q\},$$

donde Ω es el segmento en el plano complejo que une los puntos z_1 y z_2. Por eso

$$\| S \| = \sup_{x \neq 0} \frac{|(Sx, x)|}{(x, x)} = \sup_{z \in \Omega} |1 - \tau z|$$

y el parámetro τ se busca, partiendo de la condición $\min\limits_{\tau} \max\limits_{z \in \Omega} |1 - \tau z|$.

Investiguemos la función $\varphi(z) = |1 - \tau z|$. Puesto que las líneas de nivel $|1 - \tau z| = \rho_0$ son circunferencias concéntricas con centro en el punto $1/\tau$ y de radio $R = \rho_0 / |\tau|$, entonces para el valor óptimo del parámetro $\tau = \tau_0$ los puntos z_1 y z_2 deben estar sobre una línea de nivel. Por consiguiente, deben cumplirse las igualdades

$$|1 - \tau_0 z_1| = \rho_0, \quad |1 - \tau_0 z_2| = \rho_0,$$

al mismo tiempo $|1 - \tau_0 z| \leqslant \rho_0$ para $z \in \Omega$.

Escribamos estas igualdades en la forma equivalente.

$$\left| \frac{1 - \tau_0 z_2}{1 - \tau_0 z_1} \right| = 1, \quad \rho_0 = \frac{|z_2 - z_1|}{|z_1| \left| \frac{z_2}{z_1} - \frac{1 - \tau_0 z_2}{1 - \tau_0 z_1} \right|}.$$

Ya que en virtud de la primera igualdad al cambiar τ_0 el número complejo

$$z = \frac{1 - \tau_0 z_2}{1 - \tau_0 z_1}$$

recorre la circunferencia unidad en el plano complejo con centro en el origen de coordenadas, entonces ρ_0 será mínimo si se cumple la igualdad

$$\frac{1 - \tau_0 z_2}{1 - \tau_0 z_1} = -\frac{z_2}{z_1} \frac{|z_1|}{|z_2|},$$

Esta condición da el siguiente valor τ_0:

$$\tau_0 = \frac{|z_2| / z_2 + |z_1| / z_1}{|z_1| + |z_2|}. \tag{6}$$

Con esto valor $\tau = \tau_0$ para la norma del operador S es cierta la estimación

$$\| S \| = \rho_0 = \frac{|z_2 - z_1|}{|z_1| + |z_2|} \tag{7}$$

utilizando el cual, para el error x_n del esquema iterativo (2) obtendremos la estimación

$$\| x_n \|_B \leqslant \rho_0^n \| x_0 \|_B. \tag{8}$$

Hallemos ahora las condiciones bajo las cuales $\rho_0 < 1$. Puesto que es válida la desigualdad

$$|z_2 - z_1| = |z_1| \left| \frac{z_1}{|z_1|} - \frac{z_1}{|z_1|} \right| \leqslant |z_1| \left(1 + \frac{|z_2|}{|z_1|}\right) = |z_1| + |z_2|,$$

y al mismo tiempo aquí se alcanza la igualdad solamente al cumplirse la condición

$$\frac{z_1}{|z_1|} = -\frac{z_2}{|z_1|} \frac{|z_1|}{|z_2|} = -\frac{z_2}{|z_2|}, \tag{9}$$

entonces $\rho_0 < 1$, si no tiene lugar (9).

En el caso a examinar $z_1 = \gamma_1 + q$ y $z_2 = \gamma_2 + q$. De (9) encontramos fácilmente que en dos casos $\rho_0 < 1$: o bien $q_2 \neq 0$ y γ_1 y γ_2 son cualesquiera, o $q_2 = 0$, pero γ_1 y γ_2 están subordinados a la condición $(\gamma_1 + q_1)(\gamma_2 + q_1) > 0$. A continuación consideraremos que estas condiciones son cumplidas. Entonces el proceso iterativo (2) será convergente.

TEOREMA 3. *Sea A un operador hermitiano y supongamos que sean cumplidas las desigualdades (3). El proceso iterativo (2) con al parámetro*

$$\tau = \tau_0 = \frac{1}{|\gamma_1 + q| + |\gamma_2 + q|} \left(\frac{|\gamma_1 + q|}{\gamma_1 + q} + \frac{|\gamma_2 + q|}{\gamma_1 + q} \right)$$

converge en H, y para el error tiene lugar la estimación (8), donde

$$\rho_0 = \frac{|\gamma_2 - \gamma_1|}{|\gamma_1 + q| + |\gamma_2 + q|}.$$

OBSERVACION. Mas arriba fue resuelto el problema de encontrar el parámetro óptimo τ de la condición $\min\limits_{\tau} \max\limits_{z \in \Omega} |1 - \tau z|$, donde Ω es el segmento del plano complejo que une los dos puntos z_1 y z_2. Es fácil hallar la solución de este problema también en el caso, cuando Ω es un círculo con centro en el punto z_0 de radio $r_0 < |z_0|$, es decir, que no

incluye en sí el origen de coordenadas. La solución del problema planteado tiene la forma

$$\tau_0 = \frac{1}{z_0}, \quad \sup_{z \in \Omega} |1 - \tau_0 z| = \rho_0 = \frac{r_0}{|z_0|} < 1.$$

Examinemos ahora el empleo del método construido para encontrar la solución del siguiente problema de diferencias:

$$\begin{aligned} &\Lambda u - qu = -\varphi(x), \quad x \in \omega, \\ &u(x) = g(x), \quad x \in \gamma, \quad q = q_1 + iq_2, \\ &\Lambda = \Lambda_1 + \Lambda_2, \quad \Lambda_\alpha u = (a_\alpha u_{\bar{x}_\alpha})_{x_\alpha}, \quad \alpha = 1, 2, \end{aligned} \tag{10}$$

donde $\bar{\omega} = \{x_{ij} = (ih_1, jh_2) \in \bar{G}, \ 0 \leqslant i \leqslant N_1, \ 0 \leqslant j \leqslant N_2, \ h_\alpha N_\alpha = l_\alpha, \ \alpha = 1, 2\}$ es una red en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \ \alpha = 1, 2,\}$, y los coeficientes $a_\alpha(x)$ son reales y satisfacen las condiciones

$$0 < c_1 < a_\alpha(x) \leqslant c_2, \quad x \in \bar{\omega} \tag{11}$$

En el caso que se examina H es el espacio de las funciones reticulares de valores complejos, definidas sobre ω, con el producto escalar

$$(u, v) = \sum_{x \in \omega} u(x) \bar{v}(x) h_1 h_2.$$

El problema (10) se escribe en forma de la ecuación (1), donde el operador A se define de la manera usual: $Ay = -\Lambda \overset{\circ}{y}$, donde $\overset{\circ}{y} \in H$, $\overset{\circ}{y}(x) = \overset{\circ}{y}(x)$ para $x \in \omega$, $\overset{\circ}{y}(x) = 0$, $x \in \gamma$.

Para resolver la ecuación (1) construida examinemos el esquema iterativo explícito (2).

Utilizando las fórmulas de Green de diferencias para funciones de valores complejos, y también las desigualdades (11), nos cercioramos de que el operador A es hermitiano en H, y en las desigualdades (3)

$$\gamma_1 = c_1 \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha},$$

$$\gamma_2 = c_2 \sum_{\alpha=1}^{2} \frac{4}{h_\alpha^2} \cos^2 \frac{\pi h_\alpha}{2l_\alpha}.$$

Si elegimos el parámetro τ de acuerdo con el teorema 3, entonces para el error $x_n = y_n - u$ tendrá lugar la estimación (8), donde está definido en el teorema 3.

En el caso particular, cuando $l_1 = l_2 = l$, $N_1 = N_2 = N$ y $q_1 = O(1)$, $q_2 = O(1)$, obtendremos $\rho_0 = 1 - O(N^{-2})$. Por consiguiente, para alcanzar una exactitud ε dada se exige cumplir $n_0(\varepsilon) = O\left(N^2 \ln\frac{1}{\varepsilon}\right)$ iteraciones.

2. Método de las direcciones variables. Examinemos de nuevo la ecuación (1) y supongamos que el operador A se puede representar en forma de la suma de dos operadores hermitianos conmutativos A_1 y A_2:

$$A = A_1 + A_2, \quad A_1A_2 = A_2A_1, \quad A_\alpha = A_\alpha^*, \quad \alpha = 1, 2. \quad (12)$$

Sean δ y Δ las cotas de los operadores A_1 y A_2, es decir,

$$\delta E \leqslant A_\alpha \leqslant \Delta E, \quad \alpha = 1, 2. \quad (13)$$

Para resolver la ecuación (1) examinaremos el esquema iterativo de dos capas implícito (2), en el cual el operador B está prefijado de la siguiente forma:

$$B = (\omega e + A_1 + q_0 E)(\omega E + A_2 + q_0 E), \quad q_0 = 0{,}5\, q, \quad (14)$$

y los parámetros τ y ω están enlazados por la relación $\tau = 2\omega$. Nosotros obtuvimos un esquema iterativo análogo en el capítulo XI durante la construcción del método de las direcciones variables. Observemos que para encontrar y_{k+1} en el esquema (2), (14) se puede utilizar el siguiente algoritmo:

$$(\omega E + C_1)y_{k+1/2} = (\omega E - C_2)y_k + f,$$
$$(\omega E + C_2)y_{k+1} = (\omega - C_1)y_{k+1/2} + f, \quad k = 0\ 1, \ldots$$

donde para abreviar las notaciones $C_\alpha = A_\alpha + q_0 E$, $\alpha = 1, 2$.

Pasemos a investigar la convergencia del esquema (2) (14) en la norma de H. Aprovechando la conmutividad de los operadores A_1 y A_2 obtendremos la ecuación para el error z_k

$$z_{k+1} = S_1 S_2 z_k, \quad k = 0, 1, \ldots, \quad (15)$$

$$S_\alpha = (\omega E + C_\alpha)^{-1}(\omega E - C_\alpha), \quad \alpha = 1, 2, \quad (16)$$

al mismo tiempo los operadores S_1 y S_2 conmutan. De (15) hallaremos

$$z_n = S_1^n S_2^n z_0, \quad \|z_n\| \leqslant \|S_1^n\| \, \|S_2^n\| \, \|z_0\|. \quad (17)$$

Estimemos la norma del operador S_α^n, $\alpha = 1, 2$. Por cuanto C_α es un operador normal ($C_\alpha^* C_\alpha = C_\alpha C_\alpha^*$, $\alpha = 1, 2$), entonces también será normal el operador S_α. Por eso $\| S_\alpha^n \| = \| S_\alpha \|^n$ y es suficiente estimar la norma del mismo operador S_α.

Ya que la norma de un operador normal es igual a su radio espectral (véase el cap. V, § 1, punto 2), entonces de (16) obtendremos

$$\| S_\alpha \| = \max_{\lambda_\alpha} \left| \frac{\omega - \lambda_\alpha}{\omega + \lambda_\alpha} \right|, \tag{18}$$

donde los λ_α son los valores propios del operador C_α. En virtud de las suposiciones (12) y (13) hechas con respecto a los operadores A_α obtenemos, que $\lambda_\alpha \in \Omega = \{z = z_1 + a(z_2 - z_1),\ 0 \leqslant a \leqslant 1,\ z_1 = \delta + q_0,\ z_2 = \Delta + q_0\}$ para $\alpha = 1, 2$. Por lo tanto, de (18) obtenemos que

$$\| S_\alpha \| \leqslant \max_{z \in \Omega} \left| \frac{\omega - z}{\omega + z} \right|, \quad \alpha = 1, 2. \tag{19}$$

Planteemos ahora el problema de elegir el parámetro ω de la condición de mínimo del segundo miembro de la desigualdad (19).

Examinemos la aplicación fraccional lineal

$$w = (w - z)/(\omega + z), \quad \omega \neq 0, \tag{20}$$

la cual establece una relación entre los puntos del z-plano y los puntos del w-plano. De las propiedades de la transformación (20) se deduce que a las circunferencias $|\overline{w}| = \rho_0$ en el $\overline{w}$ — plano para $\rho \neq 1$ le corresponden circunferencias en el z—plano, y a la circunferencia unidad le corresponde en el z-plano una recta que pasa a través del origen de coordenadas. Los puntos de la recta indicada tienen un argumento que se diferencia del argumento de ω en $\pm\pi/2$.

Hallemos en el z-plano el centro y el radio de la circunferencia que corresponde a la circunferencia $|\overline{w}| = \rho_0 \neq 1$ en el $\overline{w}$ — plano. Para esto expresemos de (20) z mediante w

$$z = \omega(1 - w)/(1 + w),$$

y utilizando esta relación calculemos

$$\left| \frac{1 + |w|^2}{1 - |w|^2} \omega - z \right| = \left| \frac{1 + |w|^2}{1 - |w|^2} \omega - \frac{1 - w}{1 + w} \omega \right| = \frac{2|\omega|}{|1 - |w|^2|} \times \frac{|w + |w|^2|}{|1 + w|}.$$

Ya que

$$|w+|w|^2| = |w+w\overline{w}| = |w||-1+\overline{w}| = |w||1+w|,$$

entonces finalmente obtenemos

$$\left|\frac{1+|w|^2}{1-|w|^2}\right|\omega - z\Big| = \frac{2|w||\omega|}{|1-|w|^2|}.$$

De aquí se deduce que a las circunferencias $|w| = \rho_0 < 1$ corresponden las circunferencias en el z-plano con centro en el punto z_0 de radio R, donde

$$z_0 = \frac{1+\frac{a}{\omega}}{1-\rho_0^2}\omega, \quad R = \frac{2\rho_0|\omega|}{1-\rho_0^2}. \tag{21}$$

Notemos, además, que en virtud de la biunivocidad de la aplicación (20), las igualdades

$$\left|\frac{\omega - z}{\omega + z}\right| = \rho_0 < 1, \quad |z_0 - z| = R \tag{22}$$

son equivalentes

Regresemos al problema planteado. Examinemos la función

$$\varphi(z) = |w| = \left|\frac{\omega - z}{\omega + z}\right|.$$

De lo dicho más arriba se desprende que las líneas de nivel $\varphi(z) = \rho_0$ para $\rho_0 < 1$ son circunferencias con centro en el punto z_0 de radio R, donde z_0 y R están determinados en (21). Para diferentes ρ_0 estas circunferencias no se intersecan, al mismo tiempo la circunferencia, correspondiente a un valor menor de ρ_0, está contenida en el interior de una circunferencia, correspondiente a un valor mayor de ρ_0. De aquí obtenemos que para el valor óptimo $\omega = \omega_0$ los puntos z_1 y z_2 deben estar sobre una línea del nivel:

$$\left|\frac{\omega_0 - z_1}{\omega_0 + z_1}\right| = \rho_0 < 1, \quad \left|\frac{\omega_0 - z_2}{\omega_0 + z_2}\right| = \rho_0 < 1, \tag{23}$$

con esto se cumplirá la igualdad

$$\max_{z \in \Omega}\left|\frac{\omega_0 - z}{\omega_0 + z}\right| = \rho_0.$$

El parámetro ω_0 debe ser elegido de la condición de mínimo de ρ_0.

Hallemos el ω_0 óptimo y calculemos ρ_0. De (23) en virtud de (22) obtenemos

$$|z_0 - z_1| = R_0, \quad |z_0 - z_2| = R_0,$$

$$z_0 = \frac{1+\rho_0^2}{1-\rho_0^2}\omega_0, \quad R_0 = \frac{2\rho_0|\omega_0|}{1-\rho_0^2}$$

ó

$$\left|\frac{z_0-z_2}{z_0-z_1}\right| = 1, \quad \frac{2\rho_0}{1+\rho_0^2} = \frac{R_0}{z_0} = \frac{|z_2-z_1|}{|z_1|\left|\frac{z_2}{z_1} - \frac{z_0-z_2}{z_0-z_1}\right|}. \tag{24}$$

Observemos que ρ_0 es mínimo, cuando es mínimo $\frac{2\rho_0}{1+\rho_0^2}$, y esto tiene lugar, si exigimos el cumplimiento de la igualdad

$$\frac{z_0-z_2}{z_0-z_1} = -\frac{z_2}{z_1}\cdot\frac{|z_1|}{|z_2|} \tag{25}$$

Sustituyendo esta expresión en (24), obtendremos

$$\frac{2\rho_0}{1+\rho_0^2} = \frac{|z_2-z_1|}{|z_1|+|z_2|}.$$

De aquí hallaremos fácilmente

$$\gamma_1 = \frac{(1-\rho_0^2)}{1+\rho_0^2} = \frac{|z_1|+|z_2|-|z_2-z_1|}{|z_1|+|z_2|}, \tag{26}$$

$$\gamma_2 = \frac{(1+\rho_0)^2}{1+\rho_0^2} = \frac{|z_1|+|z_2|+|z_2-z_1|}{|z_1|+|z_2|},$$

$$\xi = \frac{\gamma_1}{\gamma_2} = \left(\frac{1-\rho_0}{1+\rho_0}\right)^2.$$

Por consiguiente,

$$\rho_0 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \frac{1-\rho_0^2}{1+\rho_0^2} = \sqrt{\gamma_1\gamma_2}$$

y además,

$$z_0 = \frac{1+\rho_0^2}{1-\rho_0^2}\omega_0 = \frac{\omega_0}{\sqrt{\gamma_1\gamma_1}}.$$

Sustituyendo esta expresión en (25) hallaremos el valor óptimo del parámetro ω_0:

$$\omega_0 = \frac{|z_1|+|z_2|}{|z_1|/z_1+|z_2|/z_2}\sqrt{\gamma_1\gamma_2}. \tag{27}$$

De esta forma, para el $\omega = \omega_0$ óptima está obtenida una estimación de la norma del operador S_α: $\|S_\alpha\| \leqslant \rho_0$,

$\alpha = 1, 2$. Sustituyéndolo en (17), hallaremos una estimación para el error z_n:

$$\| z_n \| \leqslant \rho_0^{2n} \| z_0 \|, \quad \rho_0 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}. \tag{28}$$

Razonando como en el método de iteración simple, hallaremos que la desigualdad $\gamma_1 > 0$, y junto con ella la desigualdad $\rho_0 < 1$ tendrán lugar en dos casos: o bien para $q_2 \neq 0$, ó para $q_2 = 0$, pero $(\delta + 0{,}5\ q_1)(\Delta + 0{,}5\ q_1) > 0$.

Así, está demostrado el siguiente

TEOREMA 4. *Sean cumplidas las condiciones (12), dadas* δ y Δ *de las desigualdades (13) y o bien* $q_2 \neq 0$, *ó* $q_2 = 0$ *y* $(\delta + 0{,}5\ q_1)(\Delta + 0{,}5\ q_1) > 0$. *Para el método de las direcciones variables (2), (14), en el cual el parámetro iterativo* $\omega = \omega_0$, *está elegido por medio de la fórmula (27), y* $\tau = 2\omega_0$ *es válida la estimación (28), donde* γ_1 y γ_2 *están definidas en (26),* $z_1 = \delta + 0{,}5\ q$ *y* $z_2 = \Delta + 0{,}5\ q$.

OBSERVACION 1. La solución del problema $\min\limits_{\omega} \max\limits_{z \in \Omega} \left| \frac{\omega - z}{\omega + z} \right|$ donde Ω es el círculo con centro en el punto z_0 del radio $r_0 < |z_0|$ tiene la forma

$$\omega_0 = z_0 \sqrt{\gamma_1 \gamma_2}, \quad \rho_0 = \max_{z \in \Omega} \left| \frac{\omega - z}{\omega + z} \right| = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2},$$

donde $\gamma_1 = 1 - r_0/|z_0|$, $\gamma_2 = 1 + r_0/|z_0|$.

OBSERVACIÓN 2. Si en lugar de las desigualdades (13) son prefijadas las desigualdades $\delta_\alpha E \leqslant A_\alpha \leqslant \Delta_\alpha E$, $\alpha = 1, 2$, entonces en el teorema 4 se debe poner $\delta = \min(\delta_1, \delta_2)$ y $\Delta = \max(\Delta_1, \Delta_2)$.

§ 3. Métodos iterativos generales para ecuaciones con un operador degenerado

1. Esquemas iterativos en el caso del operador no degenerado *B*. Sea en el espacio de Hilbert, de dimensión finita $H = H_N$ dada la ecuación

$$Au = f \tag{1}$$

con el operador degenerado A. Esto último significa que la igualdad $Au = 0$ ocurre para algún $u \neq 0$. Recordemos (véase el cap. V, § 2, punto 2) la información relacionada con el problema de la resolución de la ecuación (1).

Sea $\ker A$ el núcleo del operador A, es decir, el conjunto de los elementos $u \in H$, para los cuales $Au = 0$. Mediante $\mathrm{im}A$ — transformada del operador A — designaremos el conjunto de los elementos del tipo $y = Au$, donde $u \in H$. Es conocido que tienen lugar las siguientes descomposiciones ortogonales del espacio H en sumas directas de dos subespacios:

$$H = \ker A \oplus \mathrm{im}\, A^*, \quad H = \ker A^* \oplus \mathrm{im}\, A. \qquad (2)$$

Esto significa que cualquier elemento $u \in H$ se puede representar en la forma $u = \bar{u} + \tilde{u}$, donde $\bar{u} \in \mathrm{im}\, A^*$ y $\tilde{u} \in \ker A$, y al mismo tiempo $(\bar{u}, \tilde{u}) = 0$. Análogamente $u = \bar{u} + \tilde{u}$, donde $\bar{u} \in \mathrm{im}\, A$ y $\tilde{u} \in \ker A^*$, $(\bar{u}, \tilde{u}) = 0$.

Supongamos que en la ecuación (1) $f = \bar{f} + \tilde{f}$, donde $\bar{f} \in \mathrm{im}\, A$ y $\tilde{f} \in \ker A^*$. Se llama solución generalizada de (1) el elemento $u \in H$, para el cual $Au = \bar{f}$; ella proporciona el mínimo al funcional $\| Au - f \|$. La solución generalizada no es única y se determina con exactitud hasta un elemento de $\ker A$. La solución normal es la solución generalizada que tiene norma mínima. La solución normal es única y pertenece a $\mathrm{im}\, A^*$.

Nuestra tarea es la construcción de métodos que permitan hallar aproximadamente la solución normal de la ecuación (1). Con esto exigiremos que la solución aproximada al igual que la solución normal exacta, pertenezca al subespacio $\mathrm{im}\, A^*$.

Para la resolución del problema planteado vamos a utilizar el esquema de dos capas implícito

$$B\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \; y_0 \in H \qquad (3)$$

Primeramente estudiaremos el caso del operador B no degenerado en H. Las exigencias generales al proceso iterativo son siguientes:

a) las iteraciones se realizan según el esquema (2), la aproximación es $y_n \in \mathrm{im}\, A^*$, mientras que las aproximaciones intermedias y_k pueden pertenecer a H;

b) la estructura concreta de los subespacios $\ker A$, $\ker A^*$, $\mathrm{im}\, A$ e $\mathrm{im}\, A^*$ no se utiliza en el proceso de las iteraciones.

Hallemos las condiciones sobre el operador B, la aproximación inicial y_0 y los parámetros τ_k, $k = 1, 2, \ldots, n$,

que aseguran el cumplimiento de las exigencias formuladas más arriba.

CONDICION 1. Sea el operador B, tal que

$$Bu \in \ker A^*, \text{ si } u \in \ker A, \tag{4}$$

$$Bu \in \operatorname{im} A, \text{ si } u \in \operatorname{im} A^*. \tag{5}$$

Tiene lugar

LEMA 2. *Si para los operadores A y B son válidas las igualdades*

$$A^*B = CA, \quad BA^* = AD, \tag{6}$$

donde C y D son ciertos operadores, entonces se cumplen las condiciones (4) y (5).

En efecto, supongamos que se cumplen las igualdades (6). Si $u \in \ker A$, entonces $Au = 0$ y, por consiguiente, $A^*Bu = CAu = 0$. Por eso $Bu \in \ker A^*$, y (4) se cumple. Sea ahora $u \in \operatorname{im} A^*$, es decir $u = A^*v$, donde $v \in H$. Entonces $Bu = BA^* v = ADv \in \operatorname{im} A$. Por lo tanto, la condición (5) está cumplido. El lema está demostrado.

COROLARIO. *Para el caso $A = A^*$ las condiciones del lema 2 se cumplirán, si los operadoes A y B conmutan: $AB = BA$.*

Citemos otra serie de afirmaciones que se desprenden de (4) y (5).

LEMA 3. *Sean cumplidas las condiciones (4) y (5). Entonces*

$$B^{-1}u \in \ker A, \quad \text{si} \quad u \in \ker A^*, \tag{7}$$

$$B^{-1}u \in \operatorname{im} A^*, \quad \text{si} \quad u \in \operatorname{im} A, \tag{8}$$

y el operador AB^{-1} no es degenerado sobre im A.

En realidad, sea $u \in \ker A^*$ y $u \neq 0$. Designemos $v = B^{-1}u$, y supongamos que $v \in \operatorname{im} A^*$. Entonces en virtud de (5) $u = Bv \in \operatorname{im} A$. Pero ya que $u \neq 0$, y los subespacios im A u ker A^* son ortogonales entonces la suposición hecha no es cierta. Por consiguiente, $v = B^{-1}u \in \ker A$, y (7) está demostrado. Analogamente se demuestra (8).

Demostremos ahora la no degeneración de AB^{-1} sobre el subespacio im A. En efecto, sea $u \in \operatorname{im} A$. Entonces en virtud de (8) $B^{-1}u \in \operatorname{im} A^*$ y, por lo tanto, $B^{-1}u \perp \ker A$. De aquí obtenemos, que $AB^{-1}u \neq 0$, y por eso $(AB^{-1}u, AB^{-1}u) > 0$. El lema está demostrado.

Regresemos ahora al esquema (3) y veamos que nos da la condición 1. De acuerdo con la descomposición de H del tipo (2) representemos f y y_k para cualquier k en la forma

$$\begin{aligned} &f=\overline{f}+\widetilde{f}, \quad \overline{f}\in \operatorname{im} A, \quad \widetilde{f}\in \ker A^*, \\ &y_k=\overline{y}_k+\widetilde{y}_k, \quad \overline{y}_k\in \operatorname{im} A^*, \quad \widetilde{y}_k\in \ker A. \end{aligned} \tag{9}$$

Utilizando (9), escribiremos el esquema (3) de la siguiente forma:

$$B\frac{\overline{y}_{k+1}-\overline{y}_k}{\tau_{k+1}}+B\frac{\widetilde{y}_{k+1}-\widetilde{y}_k}{\tau_{k+1}}+A\overline{y}_k=\overline{f}+\widetilde{f}, \quad k=0,\ 1,\ \ldots \tag{10}$$

De (4) y (5) obtendremos que el primer sumando del primer miembro de (10) pertenece a im A, y el segundo a ker A^*. Por eso de (10) hallaremos la ecuación

$$B\frac{\overline{y}_{k+1}-\overline{y}_k}{\tau_{k+1}}+A\overline{y}_k=\overline{f}, \quad k=0,\ 1,\ \ldots,\ \overline{y}_0\in \operatorname{Im} A^* \tag{11}$$

para la componente $\overline{y}_k \in$ im A^* y la ecuación

$$B\frac{\widetilde{y}_{k+1}-\widetilde{y}_k}{\tau_{k+1}}=\widetilde{f}, \quad k=0,\ 1,\ \ldots,\ \widetilde{y}_0\in \ker A \tag{12}$$

para la componente $\widetilde{y}_k \in \ker A$.

Hallemos las condiciones bajo el cumplimiento de las cuales $y_n \in$ im A^*. De (9) se deduce que si $y_n = 0$, entonces $\overline{y}_n = y_n \in$ im A^*. Hallemos de (12) la expresión explícita para $\widetilde{y}_n$ e igualémosla a cero. Entonces estará cumplida la exigencia a) formulada.

De (12) obtendremos

$$\widetilde{y}_{k+1}=\widetilde{y}_k+\tau_{k+1}B^{-1}\widetilde{f}=\ldots=\widetilde{y}_0+\sum_{j=1}^{k+1}\tau_j B^{-1}\widetilde{f}.$$

De aquí se deducen

CONDICIONES 2. Sea $y_0 = A^*\varphi$, donde $\varphi \in H$, y supongamos que los parámetros τ_k, $k = 1, 2, \ldots, n$, satisfacen la exigencia

$$\sum_{j=1}^{n}\tau_j=0, \tag{13}$$

si $f \in H$. Si $f \perp \ker A^*$, entonces no se impone restricción sobre los parámetros τ_k.

Aclaremos la elección de la aproximación inicial y_0. Puesto que para cualquier $\varphi \in H$ tenemos que $y_0 = A^*\varphi \in$ $\in$ im A^*, entonces en la descomposición (9) $\tilde{y}_0 = 0$ y $\bar{y}_0 = y_0$. En particular, eligiendo $\varphi = 0$, obtendremos la aproximación inicial $y_0 = 0$.

De esta forma, si se cumplen las condiciones 2, entonces $y_n = \bar{y}_n$. Por eso el proceso iterativo (3) convergerá y dará una solución normal aproximada de la ecuación (1), si converge el proceso iterativo (11), es decir, si la sucesión $\bar{y}_k$ converge a la solución normal $\bar{u}$.

OBSERVACIÓN 1. Las condiciones 2 permiten separar de la arpoximación iterativa y_n su proyección sobre im A^* es decir, hallar $\bar{y}_n$ sin la utilización de los mismos subespacios ker A, im A, ker A^* ó im A^*. Después, si se conoce que $\sum_{j=1}^{n} \tau_j \|B^{-1}\tilde{f}\|$ es pequeña, es decir, es pequeña $\|\tilde{y}_n\|$, entonces en vritud de la igualdad $\|y_n - \bar{u}\| = \|\bar{y}_n - \bar{u}\| + \|\tilde{y}_n\|$ se puede tomar y_n en calidad de solución aproximada y renunciar la restricción (13). En este caso $y_n \in$ im A^*.

OBSERVACIÓN 2. A la condición de que todos los elementos del subespaio ker A^* son conocidos, podemos limitarnos al examen del caso $f \perp \ker A^*$, restando de f si es necesario su proyección sobre el ker A^*. Si examinamos el proceso iterativo no estacionario

$$B_{k+1}\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, y_0 = A^*\varphi,$$

y exigimos el cumplimiento de las condiciones 1, donde B está sustituido por B_k, $k = 1, 2, \ldots$, entonces todos los $y_k \in$ im A^* y no es necesario imponer ninguna restricción complementaria sobre los τ_k.

2. Método iterativo de los defectos mínimos. Examinemos ahora el problema sobre la elección de los parámetros iterativos τ_k para el esquema (3). Supondremos que el operador B satisface las condiciones 1, y que las restricciones sobre la elección de la aproximación inicial y_0 y sobre los parámetros τ_k están prefijadas por las condiciones 2.

Más arriba fue mostrado que los parámetros τ_k deben elegir de la condición de convergencia del proceso iterativo (11) a la solución normal $\bar{u}$ de la ecuación (1).

Estudiemos el esquema iterativo (11). Primeramente observemos que el operador $D = A^*A$ es definido positivo sobre im A^*. En efecto, sea $u \in \text{im } A^*$ y $u \neq 0$. Ya que $u \perp \ker A$, entonces $Au \neq 0$ y, por consiguiente, $(Du, u) = \| Au \|^2 > 0$. El operador D genera el espacio energético H_D, que consiste de los elementos de im A^*, en el cual el producto escalar se define de la manera usual: $(u, v)_D = (Du, v)$, $u \in \text{im } A^*$, $v \in \text{im } A^*$.

Planteemos ahora el problema de elegir los parámetros τ_{k+1} el esquema (11) de la condición de mínimo de $\|\bar{z}_{k+1}\|_D$, donde $\bar{z}_{k+1}$ es error: $z_{k+1} = \bar{y}_{k+1} - \bar{u}$, $A\bar{u} = \bar{f}$ y $\bar{u}$ es la solución normal de la ecuación (1)

Para el error $\bar{z}_k \in \text{im } A^*$ de (11) obtenemos la siguiente ecuación

$$\bar{z}_{k+1} = (E - \tau_{k+1} B^{-1} A)\, \bar{z}_k. \tag{14}$$

De aquí hallaremos

$$\| \bar{z}_{k+1} \|_D^2 = \| \bar{z}_k \|_D^2 - 2\tau_{k+1} (AB^{-1}A\bar{z}_k, A\bar{z}_k) + \tau_{k+1}^2 \| AB^{-1}Az \|^2.$$

Notemos que en virtud del lema 3 $\|AB^{-1}A\bar{z}_k\| > 0$, $(A\bar{z}_k \in \in \text{im } A)$. $(Az_k \in \text{im } A)$. Por eso el mínimo de $\|\bar{z}_{k+1}\|_D^2$ se alcanza para

$$\tau_{k+1} = \frac{(AB^{-1}A\bar{z}_k, A\bar{z}_k)}{(AB^{-1}A\bar{z}_k, AB^{-1}A\bar{z}_k)} \tag{15}$$

y es igual a

$$\| \bar{z}_{k+1} \|_D^2 = \rho_{k+1}^2 \| \bar{z}_k \|_D^2,$$

$$\rho_{k+1}^2 = 1 - \frac{(AB^{-1}A\bar{z}_k, A\bar{z}_k)_2}{\| AB^{-1}A\bar{z}_k \|^2 \| A\bar{z}_k \|^2}. \tag{16}$$

La fórmula (15 aún no está apta para los cálculos, ya que contiene magnitudes incógnitas. Transformemos dicha fórmula. Utilizando la descomposición (9), obtendremos

$$A\bar{z}_k = A\bar{y}_k - \bar{f} = Ay_k - \bar{f} = r_k + \tilde{f}, \tag{17}$$

donde $r_k = Ay_k = f$ es el defecto. Ya que $\tilde{f} \in \ker A^*$, entonces en virtud del lema 3 $B^{-1}\tilde{f} \in \ker A$ y, por consiguiente, $AB^{-1}A\bar{z}_k = AB^{-1}r_k$. Sustituyendo esta expresión, y tam-

bién (17) en (15) y teniendo en cuenta la igualdad $A^*\tilde{f} = 0$, obtendremos

$$\tau_{k+1} = \frac{(AB^{-1}r_k, r_k)}{(AB^{-1}r_k, AB^{-1}r_k)} = \frac{(Aw_k, r_k)}{(Aw_k, Aw_k)}, \tag{18}$$

donde la corrección w_k se encuentra de la ecuación $Bw_k = r_k$.

Notemos, que (18) coincide con la fórmula para el parámetro iterativo τ_{k+1} del método de los defectos mínimos, examinado en el capítulo VIII para una ecuación con el operador A no degenerado.

Obtengamos ahora la estimación de la velocidad de convergencia para el método construido. Multipliquemos (14) a la izquierda por A, calculemos la norma de los primero y segundo miembros y, teniendo en cuenta que $\| A\bar{z}_k \| = \| \bar{z}_k \|_D$ obtendremos la siguiente estimación:

$$\| \bar{z}_{k+1} \|_D \leqslant \| E - \tau_{k+1} AB^{-1} \|_{\mathrm{im}A} \| \bar{z}_k \|_D \tag{19}$$

para cualquier τ_{k+1}. De (16) y (19) obtenemos para cualquier τ_{k+1}

$$\rho_{k+1} \leqslant \| E - \tau_{k+1} AB^{-1} \|_{\mathrm{im}A}. \tag{20}$$

Si designamos

$$\rho_0 = \min_{\tau} \| E - \tau AB^{-1} \|_{\mathrm{im}A},$$

entonces de (16) y (20) se deducirá la estimación para el error

$$\| \bar{z}_{k+1} \|_D \leqslant \rho_0 \| \bar{z}_k \|_D. \tag{21}$$

Aquí la notación $\| S \|_{\mathrm{im}A}$ se utiliza para designar la norma del operador S en el subespacio im A.

Si $\rho_0 < 1$, entonces el método iterativo (11), (18) convergerá en H_D y de (21) obtenemos, que

$$\| \bar{z}_k \|_D \leqslant \rho_0^k \| \bar{z}_0 \|_D, \quad k = 0, 1, \ldots \tag{22}$$

Nos queda solamente subordinar la elección de los parámetros τ_k a la condición (13), si $\tilde{f} \neq 0$. Para esto procederemos de la siguiente manera. Cumplamos la $(n-1)$ iteración según el esquema (3), eligiendo $y_0 = A^*\varphi$, donde $\varphi \in H$, utilizando para los parámetros τ_{k+1}, $k = 0, 1, \ldots, n-2$ la fórmula (18). Cumpliremos más otra iteración, eligiendo

$$\tau_n = -\sum_{j=1}^{n-1} \tau_j.$$

Entonces la condición (13) será cumplida y, por consiguiente, $y_n = \bar{y}_n$. Estimemos ahora la norma del error $z_n = y_n - \bar{u}$ en H_D. Ya que $y_n = \bar{y}_n$, entonces de (11) obtendremos

$$y_n = \bar{y}_{n-1} - \tau_n B^{-1}(A\bar{y}_{n-1} - \bar{f}) = \bar{y}_{n-1} - \tau_n B^{-1} A\bar{z}_{n-1}.$$

De aquí

$$z_n = \bar{z}_{n-1} - \tau_n B^{-1} A\bar{z}_{n-1}$$

y después de multiplicar por A obtendremos

$$Az_n = (E - \tau_n AB^{-1}) A\bar{z}_{n-1}.$$

Calculando la norma, obtendremos la estimación

$$\| z_n \|_D \leqslant \| E - \tau_n AB^{-1} \|_{\mathrm{im}A} \| \bar{z}_{n-1} \|_D.$$

Sustituyendo aquí (22) y teniendo en cuenta que en virtud de la elección de y_0 tenemos la igualdad $\bar{y}_0 = y_0$, hallaremos

$$\| y_n - \bar{u} \|_D \leqslant \| E - \tau_n AB^{-1} \|_{\mathrm{im}A} \rho_0^{n-1} \| y_0 - \bar{u} \|_D. \qquad (23)$$

Examinemos unos casos particulares.

1) Sea $B = E$, y el operador A autoconjugado en H. Sean γ_1 y γ_2 las constantes en las desigualdades

$$\gamma_1 (x, x) \leqslant (Ax, x) \leqslant \gamma_2 (x, x), \quad \gamma_1 > 0, \quad Ax \neq 0. \qquad (24)$$

En este caso se cumplen las condiciones 1.

Hallemos ρ_0 y estimemos la norma del operador en (23). Puesto que el operador A es autoconjugado en H, entonces, utilizando (24), obtenemos

$$\| E - \tau A \|_{\mathrm{im}A} = \sup_{Au \neq 0} \left| 1 - \tau \frac{(Au, u)}{(u, u)} \right| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |1 - \tau t|.$$

Con el cálculo del máximo indicado y la elección de τ de la condición de su mínimo nos encontramos en el capítulo VI al examinar el método de iteración simple. Allí fue hallado que

$$\min_{\tau} \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |1 - \tau t| = \rho_0 = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Así, hemos hallado ρ_0. A continuación, para $B = E$ la fórmula (15) para el parámetro τ_{k+1} se escribe en la forma

$$\tau_{k+1} = \frac{(Ax, x)}{(Ax, Ax)}, \quad x = A\bar{z}_k \in \mathrm{im}\, A.$$

Ya que $A = A^*$ y $\gamma_1 > 0$, entonces las desigualdades (24) son equivalentes a las desigualdades siguientes (véase el cap. V, el § 1, punto 3).

$$\gamma_1 (Ax, x) \leqslant (Ax, Ax) \leqslant \gamma_2 (Ax, x), \quad Ax \neq 0.$$

Por eso los parámetros τ_k para $k \leqslant n-1$ satisfacen las desigualdades $1/\gamma_2 \leqslant \tau_k \leqslant 1/\gamma_1$. De aquí hallaremos la estimación

$$0 < -\tau_n = \sum_{j=1}^{n-1} \tau_j \leqslant \frac{n-1}{\gamma_1}. \tag{25}$$

Estimemos la norma del operador en (23). Teniendo en cuenta (24) y (25), obtendremos

$$\| E - \tau_n A \|_{\mathrm{im}\, A} \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |1 - \tau_n t| =$$

$$= 1 - \tau_n \gamma_2 \leqslant 1 + (n-1)\frac{\gamma_2}{\gamma_1} = 1 + (n-1)\frac{1+\rho_0}{1-\rho_0}.$$

Sustituyamos esta estimación en (23) y hallaremos

$$\| y_n - \bar{u} \|_D \leqslant \rho_0^{n-1} \left[1 + (n-1)\frac{1+\rho_0}{1-\rho_0} \right] \| y_0 - \bar{u} \|_D. \tag{26}$$

2) Sea $B = B^*$, $A = A^*$ y $AB = BA$. Sean γ_1 y γ_2 las constantes en las desigualdades

$$\gamma_1 (Bx, x) \leqslant (Ax, x) \leqslant \gamma_2 (Bx, x), \quad \gamma_1 > 0, \; Ax \neq 0. \tag{27}$$

En este caso se cumplen las condiciones 1, el operador AB^{-1} es autoconjugado en H y se puede mostrar, que para el error del método (3), (18) será cierta la estimación (26).

3) Sean los operadores B^*A y A^*B autoconjugados en H, y γ_1 y γ_2 las constantes en (27). En este caso en virtud del lema 2 se cumplen las condiciones 1. Además, el operador AB^{-1} será autoconjugado en H. Se puede mostrar, que en este caso tiene lugar la estimación (26).

3. El método con los parámetros de Chebishev. Examinemos ahora los métodos iterativos (3), para los cuales los parámetros τ_k se eligen con la utilización de la información apriorística sobre los operadores A y B.

Primeramente citemos algunas afirmaciones auxiliares que nos serán necesarias para la exposición ulterior.

LEMA 4. *Sean cumplidas las condiciones*

$$A = A^* \geqslant 0, \quad B = B^* \geqslant 0, \quad AB = BA \tag{28}$$

y sean prefijadas las constantes γ_1 *y* γ_2 *en las desigualdades*

$$\gamma_1 (Bx, x) \leqslant (Ax, x) \leqslant \gamma_2 (Bx, x), \quad \gamma_1 > 0, \quad Ax \neq 0. \qquad (29)$$

Designemos mediante D *uno de los operadores* A, B *ó* $AB^{-1}A$ *y definamos sobre el subespacio* im A *al operador* C

$$C = D^{-1/2} (DB^{-1}A) D^{-1/2}.$$

El operador C *es autoconjugado en* im A *y satisface las desigualdades*

$$0 < \gamma_1 (x, x) \leqslant (Cx, x) \leqslant \gamma_2 (x, x), \quad x \in \text{im } A. \qquad (30)$$

Realmente de (28) y del corolario al lema 2 se desprende el cumplimiento de las condiciones 1. Además, el operador D es autoconjugado en H y definido positivo sobre im A. Para ejemplo demostremos la definición positiva del operador $D = AB^{-1}A$. Sea $u \in \text{im } A$ y $u \neq 0$. Ya que $(Du, u) = (B^{-1}Au, Au)$, y el operador B^{-1} es definido positivo en virtud de la acotación y definición positiva del operador B, entonces $(Du, u) \geqslant 0$, al mismo tiempo la igualdad a cero es posible solamente al cumplirse la condición $Au = 0$. Pero esto contradice a las suposiciones hechas.

El operador D aplica im A sobre im A, por eso existe $D^{-1/2}$, el cual también aplica este subespacio sobre sí mismo. Por consiguiente, sobre im A se puede definir el operador C indicado en el lema. El paso de (29) a (30) se demuestra igual que fue hecho en el cap. VI, § 2, punto 3. El lema está demostrado.

LEMA 5. *Sean cumplidas las condiciones*

$$B^*A = A^*B, \quad AB^* = BA^* \qquad (31)$$

y dadas γ_1 y γ_2 *en (29). Designemos* $C_1 = AB^{-1}$ *y* $C_2 = B^{-1}A$. *Los operadores* C_1 *y* C_2 *son autoconjugados en* H *y satisfacen las desigualdades*

$$\gamma_1 (x, x) \leqslant (C_1 x, x) \leqslant \gamma_2 (x, x), \quad \gamma_1 > 0, \; x \in \text{im } A, \qquad (32)$$

$$\gamma_1 (x, x) \leqslant (C_2 x, x) \leqslant \gamma_2 (x, x), \quad \gamma_1 > 0, \quad x \in \text{im } A^* \qquad (33)$$

La autoconjugación de los operadores C_1 y C_2 se deduce directamente de (31). Demostremos para ejemplo (32). Examinemos el problema en valores propios

$$AB^{-1}v - \lambda v = 0, \quad v \in H. \qquad (34)$$

Ya que el operador AB^{-1} es autoconjugado en H, entonces existe un sistema ortonormalizado de funciones propias del

problema (34) $\{v_1, v_2, \ldots, v_p, v_{p+1}, \ldots, v_N\}$. Sean $v_1, \ldots, v_p$ las funciones, correspondientes al valor propio $\lambda = 0$, mientras $v_{p+1}, \ldots, v_N$ corresponden a los λ no nulos. Es fácil ver que $v_i \in \ker A^*$, $1 \leqslant i \leqslant p$, $v_l \in \operatorname{im} A$, $p+1 \leqslant i \leqslant N$, y en virtud de la descomposición de H en subespacios (2), las funciones $v_{p+1}, \ldots, v_N$ constituyen una base en $\operatorname{im} A$. Entonces para $x \in \operatorname{im} A$ tenemos

$$x = \sum_{k=p+1}^{N} a_k v_k, \quad C_1 x = \sum_{k=p+1}^{N} \lambda_k a_k v_k,$$

y en virtud de la ortogonalidad de las funciones propias

$$(x, x) = \sum_{k=p+1}^{N} = a_k^2, \quad (C_1 x, x) = \sum_{k=p+1}^{N} \lambda_k a_k^2.$$

De aquí obtenemos las desigualdades

$$\min_{p+1 \leqslant k \leqslant N} \lambda_k (x, x) \leqslant (C_1 x, x) \leqslant \max_{p+1 \leqslant k \leqslant N} \lambda_k (x, x).$$

Quedan por hallar los valores propios mínimo y máximo, correspondientes a las funciones propias del problema (34), pertenecientes a $\operatorname{im} A$. Escribamos (34) en la forma

$$Au_k - \lambda_k B u_k = 0, \quad p+1 \leqslant k \leqslant N, \tag{35}$$

donde $u_k = B^{-1} v_k \in \operatorname{im} A^*$ y, por consiguiente, $Au_k \neq 0$. Multiplicando (35) escalarmente por u_k y utilizando (29), obtendremos, que

$$\min_{p+1 \leqslant k \leqslant N} \lambda_k = \gamma_1, \qquad \max_{p+1 \leqslant k \leqslant N} \lambda_k = \gamma_2.$$

Las desigualdades (32) están demostradas. La justeza de (33) se establece análogamente. El lema está demostrado.

Volvamos ahora al problema de elegir los parámetros iterativos para el esquema (3). Teniendo en cuenta las condiciones 2 escribámoslo en la forma siguiente:

$$B \frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = f, \quad y_0 \in A^* \varphi, \quad \sum_{j=1}^{n} \tau_j = 0. \tag{36}$$

Si se cumplen las condiciones 2, entonces es necesario elegir los parámetros τ_k de la condición de convergencia del esquema (11) bajo la restricción indicada más arriba sobre la suma de los τ_j.

Examinemos la ecuación (14) para el error del esquema (11) Si son cumplidas las condiciones del lema 4, entonces,

poniendo $\bar{z}_k = D^{-1/2} x_k$ en (14), donde D es uno de los operadores del lema 4, obtendremos la siguiente ecuación para el error equivalente:

$$x_{k+1} = (E - \tau_{k+1} C)\, x_k, \quad k = 0, 1, \ldots, x_k \in \operatorname{im} A. \quad (37)$$

El operador C también está definido en el lema 4.

Si se cumplen las condiciones del lema 5, entonces designando $B\bar{z}_k = x_k$ ó $A\bar{z}_k = x_k$, obtenemos la ecuación

$$x_{k+1} = (E - \tau_{k+1} C_1)\, x_k, \quad k = 0, 1, \ldots, x_k \in \operatorname{im} A. \quad (38)$$

En este caso $\| x_k \| = \| \bar{z}_k \|_D$, donde $D = B^*B$ ó A^*A. Si designamos $\bar{z}_k = x_k$, entonces obtendremos la ecuación

$$x_{k+1} = (E - \tau_{k+1} C_2)\, x_k, \quad k = 0, 1, \ldots, x_k \in \operatorname{im} A^*, \quad (39)$$

y en este caso $\| x_k \| = \| \bar{z}_k \|_D$, donde $D = E$. Los operadores C_1 y C_2 están definidos en el lema 5.

Así pues, en todos los casos examinados nosotros obtuvimos una ecuación del tipo

$$x_{k+1} = (E - \tau_{k+1} C)\, x_k, \quad k = 0, 1, \ldots, x_k \in H_1 \quad (40)$$

en un subespacio H_1, al mismo tiempo en virtud de los lemas 4 y 5 el operador C es autoconjugado en H_1 actúa en H_1 y satisface las desigualdades

$$\gamma_1 (x, x) \leqslant (Cx, x) \leqslant \gamma_2 (x, x), \quad \gamma_1 > 0, \quad x \in H_1, \quad (41)$$

donde γ_1 y γ_2 se toman de las desigualdades (29).

De (40) hallaremos

$$x_n = \prod_{j=1}^{n} (E - \tau_j C)\, x_0, \quad (42)$$

$$\| x_n \| \leqslant \| P_n (C) \| \, \| x_0 \|, \quad P_n (C) = \prod_{j=1}^{n} (E - \tau_j C).$$

Teniendo en cuenta la autoconjugación de C y las desigualdades (41), obtenemos

$$\| P_n (C) \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} | P_n (t) |.$$

Es facil ver que

$$\sum_{j=1}^{n} \tau_j = - P_n' (0),$$

por eso el polinomio $P_n(t)$ está normado por dos condiciones

$$P_n(0) = 1, \quad P_n'(0) = 0. \tag{43}$$

Por consiguiente, nosotros llegamos al problema de construir un polinomio de grado n, que satisfaga las condiciones (43) y con la menor desviación del cero sobre el segmento $0 < \gamma_1 \leqslant t \leqslant \gamma_2$. La construcción de este polinomio resuelve completamente el problema de la elección de los parámetros iterativos τ_k para el esquema (3).

La solución exacta de este problema no es conocida y nosotros citaremos una solución diferente del problema. Como en el método de los defectos mínimos examinado más arriba, dejaremos arbitrariedad en la elección de los parámetros $\tau_1, \tau_2, \ldots, \tau_{n-1}$, y satisfaceremos la condición $\sum_{j=1}^{n} \tau_j = 0$ a cuenta de la elección de τ_n por la fórmula

$$\tau_n = -\sum_{j=1}^{n-1} \tau_j.$$

De (42) obtendremos la siguiente estimación:

$$\| x_n \| \leqslant \| P_{n-1}(C) \| \, \| E - \tau_n C \| \, \| x_0 \|,$$

$$P_{n-1}(C) = \prod_{j=1}^{n-1} (E - \tau_j C). \tag{44}$$

Elijamos ahora los parámetros $\tau_1, \tau_2, \ldots, \tau_{n-1}$ de la condición de mínimo de la norma del polinomio operacional $P_{n-1}(C)$. Ya que sobre $P_{n-1}(C)$ no se imponen ningunas restricciones complementarias, entonces la solución del problema planteado tiene la forma (véase el cap. VI, § 2):

$$\tau_k = \frac{\tau_0}{1 + \rho_0 \mu_k}, \quad \mu_k \in \mathfrak{M}_{n-1} =$$

$$= \left\{ \cos \frac{(2i-1)\pi}{2(n-1)}, \; 1 \leqslant i \leqslant n-1 \right\}, \tag{45}$$

$k = 1, 2, \ldots, n-1$ donde son habituales las notaciones

$$\tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Con esto

$$P_{n-1}(t) = q_{n-1} T_{n-1}\left(\frac{(1 - \tau_0 t)}{\rho_0} \right), \quad \| P_{n-1}(C) \| \leqslant q_{n-1}, \tag{46}$$

donde $T_{n-1}(x)$ es el polinomio de Chebishev de primer género y de grado $n-1$,

$$q_k = 2\rho_1^k/(1+\rho_1^{2k}), \quad \rho_1 = (1-\sqrt{\xi})/(1+\sqrt{\xi}).$$

Queda por hallar una expresión explícita para τ_n. De (46) hallaremos

$$\tau_n = -\sum_{j=1}^{n-1} \tau_j = P'_{n-1}(0) = -\frac{(n-1)\tau_0}{\rho_0} q_{n-1} U_{n-2}\left(\frac{1}{\rho_0}\right), \quad (47)$$

donde $U_{n-2}(x)$ es el polinomio de Chebishev de segundo género y de grado $n-2$. Aquí fue utilizada la relación $T'_m(x) = mU_{m-1}(x)$. Calculemos $U_{n-2}(1/\rho_0)$. Ya que $\rho_0 < 1$, entonces de la forma explícita para $U_{n-2}(x)$ (véase el cap. I, § 4, punto 2):

$$U_{n-2}(x) = \frac{(x+\sqrt{x^2-1})^{n-1} - (x+\sqrt{x^2-1})^{-(n-1)}}{2\sqrt{x^2-1}}, \quad |x| \geqslant 1,$$

obtenemos como resultado de transformaciones no complejas

$$U_{n-2}\left(\frac{(1)}{\rho_0}\right) = \frac{1-\rho_1^{2(n-1)}}{2\rho_1^{n-1}} \frac{\rho_0}{\sqrt{1-\rho_0^2}}.$$

Sustituyamos esta expresión en (47) y hallemos

$$\tau_n = -\frac{(n-1)\tau_0}{\sqrt{1-\rho_0^2}} \frac{1-\rho_1^{2(n-1)}}{1+\rho_1^{2(n-1)}}. \quad (48)$$

Teniendo en cuenta la autoconjugación de C y las desigualdades (41), la fórmula (48) y la igualdad $(\tau_0\gamma_2) = 1 + \rho_0$, obtendremos

$$\| E - \tau_n C \| \leqslant \max_{\gamma_1 \leqslant t \leqslant \gamma_2} |1 - \tau_n t| = 1 - \tau_n \gamma_2 =$$

$$= 1 + (n-1)\sqrt{\frac{1+\rho_0}{1-\rho_0}} \frac{1-\rho_1^{2(n-1)}}{1+\rho_1^{2(n-1)}} = 1 + (n-1)\sqrt{\frac{1+\rho_0}{1-\rho_0}}. \quad (49)$$

Sustituyendo (49) y (46) en (44), obtenemos la siguiente estimación para la norma del error equivalente x_n:

$$\| x_n \| \leqslant \left(1 + (n-1)\sqrt{\frac{1+\rho_0}{1-\rho_0}}\right) q_{n-1} \| x_0 \|$$

bajo la condición de que los parámetros $\tau_1, \tau_2, \ldots, \tau_n$ estén elegidos por las fórmulas (45) y (48).

TEOREMA 5. *Sean los parámetros iterativos* τ_k, $k = 1, \ldots$ $\ldots, n$, *para el esquema* (3) *elegidos según las fórmulas* (45) *y* (48) *y* $y_0 = A^*\varphi$. *Entonces para el error es cierta la estimación*

$$\| y_n - \bar{u} \|_D \leqslant \left(1 + (n-1)\sqrt{\frac{1+\rho_0}{1-\rho_0}}\right) q_{n-1} \| y_0 - \bar{u} \|_D,$$

donde u *es la solución normal de la ecuación* (1), *y* D *se define de la siguiente forma:* $D = A$, B *ó* $AB^{-1}A$, *si se cumplen las condiciones del lema* 4; $D = B^*B$, A^*A *ó* E, *si se satisfacen las condiciones del lema* 5. *La información apriorística para el método con los parámetros de Chebishev son las constantes* γ_1 *y* γ_2 *de las desigualdades* (29).

§ 4. Métodos especiales

1. Problema de Neumann de diferencias para la ecuación de Poisson en un rectángulo. En el ejemplo del problema indicado ilustraremos el empleo de un esquema iterativo del operador variable B_k en la resolución de una ecuación del operador degenerado A.

Supongamos que en el rectángulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \alpha = 1, 2\}$ se exige hallar la solución de la ecuación de Poisson

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -\varphi(x), \quad x \in G, \tag{1}$$

que satisface las siguientes condiciones de contorno:

$$\begin{aligned} &\frac{\partial u}{\partial x_\alpha} = -g_{-\alpha}(x_\beta), \quad x_\alpha = 0, \quad \beta = 3 - \alpha, \\ &-\frac{\partial u}{\partial x_\alpha} = -g_{+\alpha}(x_\beta), \quad x_\alpha = l_\alpha, \quad \alpha = 1, 2. \end{aligned} \tag{2}$$

Sobre la red rectangular $\bar{\omega} = \{\bar{x_{ij}} = (ih_1\ jh_2) \in \bar{G},\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2,\ h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2\}$ al problema (1), (2) le corresponde el siguiente problema de diferencias:

$$\Lambda y = -f(x), \quad x \in \bar{\omega}, \tag{3}$$

$$\Lambda = \Lambda_1 + \Lambda_2, \quad f(x) = \varphi(x) + \frac{2}{h_1}\varphi_1(x) + \frac{2}{h_2}\varphi_2(x),$$

donde

$$\Lambda_\alpha y = \begin{cases} \frac{2}{h_\alpha} y_{x_\alpha}, & x_\alpha = 0, \\ y_{\bar{x}_\alpha x_\alpha}, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ -\frac{2}{h_\alpha} y_{\bar{x}_\alpha}, & x_\alpha = l_\alpha, \end{cases} \tag{4}$$

$$\varphi_\alpha(x) = \begin{cases} g_{-\alpha}(x_\beta), & x_\alpha = 0, \\ 0, \; h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ g_{+\alpha}(x_\beta), & x_\beta = l_\alpha. \end{cases}$$

El espacio H consiste de las funciones reticulares definidas sobre la red $\bar{\omega}$, con el producto escalar $(u, v) = \sum_{x \in \bar{\omega}} u(x) v(x) \hbar_1(x)_1 \hbar_2(x_2)$, donde $\hbar_\alpha(x_\alpha)$ es el paso medio,

$$\hbar_\alpha(x_\alpha) = \begin{cases} h_\alpha, \; h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ 0{,}5h_\alpha, \quad x_\alpha = 0, \quad l_\alpha, \; \alpha = 1, 2. \end{cases}$$

Definiremos al operador A como la suma de los operadores A_1 y A_2, donde $A_\alpha = -\Lambda_\alpha$, $\alpha = 1, 2$. Entonces el problema (3) se puede escribir en forma de la ecuación operacional

$$Au = f \tag{5}$$

con el operador A indicado.

Señalemos las siguientes propiedades de los operadores A_1 y A_2. Los operadores A_1 y A_2 son autoconjugados en H y conmutan, es decir

$$A_\alpha = A_\alpha^*, \quad \alpha = 1, 2, \quad A_1 A_2 = A_2 A_1.$$

Estas propiedades permiten, utilizando el método de separación de variables, resolver el problema en valores propios para el operador A: $Au = \lambda u$. Actuando por analogía con el caso del problema de Dirichlet, examinado detalladamente en el punto 1, del § 2, del cap. IV, obtendremos la solución del problema en la forma

$$\lambda_{k_1 k_2} = \lambda_{k_1}^1 + \lambda_{k_2}^{(2)}, \quad \lambda_{k_\alpha}^{(\alpha)} = \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{k_\alpha \pi}{2N_\alpha}, \; k_\alpha = 0, 1, \ldots, N_\alpha,$$

$$\mu_{k_1 k_2}(i, j) = \mu_{k_1}^{(1)}(i) \mu_{k_2}^{(2)}(j), \; 0 \leqslant k_\alpha \leqslant N_\alpha, \quad \alpha = 1, 2,$$

$$\mu_{k_\alpha}^{\alpha}(m) = \begin{cases} \sqrt{\frac{2}{l_\alpha}} \cos \frac{k_\alpha \pi m}{N_\alpha}, & 1 \leqslant k_\alpha \leqslant N_\alpha - 1, \\ \sqrt{\frac{1}{l_\alpha}} \cos \frac{k_\alpha \pi m}{N_\alpha}, & k_\alpha = 0, \quad N_\alpha, \; \alpha = 1, 2. \end{cases}$$

Con esto tenemos

$$A_\alpha \mu_{k_1 k_2} = \lambda_{k_\alpha}^{(\alpha)} \mu_{k_1 k_2}, \quad A\mu_{k_1 k_2} = \lambda_{k_1 k_2} \mu_{k_1 k_2}.$$

De aquí se deduce que el operador A tiene un valor propio simple igual a cero, al cual le corresponde la función propia $\mu_{00}(i, j) \equiv 1/\sqrt{l_1 l_2}$. Esta función forma la base en el subespacio ker A. Las funciones $\mu_{k_1 k_2}(i, j)$ para $0 \leqslant \leqslant k_\alpha \leqslant N_\alpha$ y $k_1 + k_2 \neq 0$ forman la base en el subespacio im A.

Para resolver la ecuación (5) examinaremos el esquema iterativo del método de las direcciones variables

$$B_{k+1} \frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \; y_0 \in H,$$
$$B_k = (\omega_k^{(1)} E + A_1)(\omega_k^{(2)} E + A_2), \quad \tau_k = \omega_k^{(1)} + \omega_k^{(2)}. \tag{6}$$

Para no imponer restricciones complementarias sobre los $\{\tau_k\}$ y los operadores $\{B_k\}$, relacionadas con la separación de la componente $y_n \in \text{im } A$, exigiremos, que el segundo miembro f sea ortogonal al ker A. Si la f dada no satisface esta condición, entonces la cambiaremos en (6) por $f_1 = f - (f, \mu_{00})\mu_{00}$.

Notemos, que los operadores B_k y A son conmutables para cualquier k. Por eso en virtud del corolario al lema 2 se cumplirán las condiciones 1 (en ellas es necesario sustituir B por el operador B_k). Además en virtud del lema 3 el operador B_k^{-1} aplica im A sobre im A.

Aprovechemos los hechos establecidos para estudiar la convergencia del esquema (6) ya que $f \in \text{im } A$, entonces, suponiendo que $y_k \in \text{im } A$, obtendremos de (6), que

$$y_{k+1} = y_k - \tau_{k+1} B_{k+1}^{-1} (Ay_k - f) \in \text{im } A.$$

Por eso, si elegimos $y_0 = 0$, entonces $y_0 \in \text{im } A$, por consiguiente, para cualquier $k \geqslant 0$ las aproximaciones iterativas $y_k \in \text{im } A$. Por consiguiente, el esquema (6) puede examinarse solamente sobre el subespacio im A.

Investiguemos la convergencia del esquema (6) en im A según la norma del espacio H_D, donde en calidad de D se puede tomar uno de los operadores E, A ó A^2. Cada uno de estos operadores será definido positivo en im A. El procedimiento para estudiar la convergencia del esquema (6) es exactamente el mismo que fue utilizado en el capítulo XI al construir el método de las direcciones variables en el caso no degenerado. Por lo tanto nos limitaremos sólo a la

formulación del problema sobre la mejor elección de los parámetros, omitiendo todos los razonamientos necesarios para esto.

Los mejores parámetros $\omega_j^{(1)}$ y $\omega_j^{(2)}$ para el esquema (6) deben ser elegidos de la condición

$$\min_{\{\omega_j\}} \max_{x,\, y \in \Omega} \left| \prod_{j=1}^{n} \frac{\omega_j^{(2)} - x}{\omega_j^{(1)} + x} \, \frac{\omega_j^{(1)} - y}{\omega_j^{(2)} + y} \right| = \rho_n,$$

donde $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$, $\Omega_1 = \{\lambda_1^{(1)} \leqslant x \leqslant \lambda_{N_1}^{(1)},\ \lambda_1^{(2)} \leqslant y \leqslant \lambda_{N_2}^{(2)}\}$, $\Omega_2 = \{\lambda_1^{(1)} \leqslant x \leqslant \lambda_{N_2}^{(1)},\ y = 0\}$ y $\Omega_3 = \{x = 0,\ \lambda_1^{(2)} \leqslant y \leqslant \lambda_{N_2}^{2}\}$.

Con esto para el error $z_n = y_n - \bar{u}$, donde $\bar{u}$ es la solución normal de la ecuación (5), será cierta la estimación

$$\| z_n \|_D \leqslant \rho_n \| z_0 \|_D.$$

Señalemos que para el ejemplo a examinar la condición de ortogonalidad de $\bar{u}$ al ker A se escribe en la forma $(u, 1) = 0$. Cualquier otra solución de la ecuación (5) se diferencia de la solución normal $\bar{u}$ en una función, igual a constante sobre la red $\bar{\omega}$. Por eso una de las posibles soluciones del problema (3) se puede separar, fijando el valor de esta solución en un nodo de la red $\bar{\omega}$.

El problema formulado más arriba para los parámetros se diferencia del examinado por nosotros en el § 1 del cap. XI, pero puede ser reducido a él mediante algunas simplificaciones y al precio de disminuir la posible velocidad de convergencia del método iterativo. Designemos

$$\delta = \min_{\alpha} \lambda_1^{(\alpha)}, \quad \Delta = \max_{\alpha} \lambda_{N_\alpha}^{\alpha}, \quad \eta = \frac{\delta}{\Delta},$$

$$\varkappa_j = \frac{1}{\Delta} \omega_j^{(1)} = \frac{1}{\Delta} \omega_j^{(2)}, \quad j = 1, 2, \ldots, n.$$

Utilizando estas notaciones y la estructura de la región Ω, el problema de la elección de los parámetros se puede formular así: elegir $\varkappa_j$, $1 \leqslant j \leqslant n$ de la condición

$$\min_{\{\varkappa_j\}} \max_{\eta \leqslant u \leqslant 1} |r_n(u, \varkappa)| = \bar{\rho}_n, \quad r_n(u, \varkappa) = \prod_{j=1}^{n} \frac{\varkappa_j - u}{\varkappa_j + u}.$$

Con esto, obviamente, $\bar{\rho}_n > \rho_n$.

Precisamente este problema fue examinado en el § 1 del cap. XI. Recordemos que allí fueron obtenidas las fór-

mulas para las $\varkappa_j$ y para el número de iteraciones $n = n_0(\varepsilon)$, los cuales garantizaban el cumplimiento de la desigualdad $\bar{\rho}_n^2 \leqslant \varepsilon$. Puesto que aquí necesitamos asegurar la estimación $\bar{\rho}_n \leqslant \varepsilon$, entonces en las fórmulas para $\varkappa_j$ y $n_0(\varepsilon)$ del cap. XI se debe poner ε^2 en lugar de ε. Entonces para el error del método (6) se cumplirá la estimación $\| z_n \|_D \leqslant \varepsilon \| z_0 \|_D$. Mostremos el aspecto de la estimación para el número de iteraciones: $n \geqslant n_0(\varepsilon)$, $n_0(\varepsilon) = \frac{1}{\pi^2} \ln \frac{4}{\eta} \ln \frac{4}{\varepsilon^2}$. Para ejemplo, si $l_1 = l_2 = l$ y $h_1 = h_2 = h$, entonces

$$\delta = \frac{4}{h^2} \operatorname{sen}^2 \frac{\pi h}{2l}, \quad \Delta = \frac{4}{h^2}, \quad \eta = \operatorname{sen}^2 \frac{\pi h}{2},$$

$$n_0(\varepsilon) = O(\ln h \ln \varepsilon).$$

Por consiguiente, para el problema de Neumann el método de direcciones variables, aún teniendo la estimación del mismo orden del número de iteraciones que para el caso del problema de Dirichlet, exige prácticamente en dos veces más iteraciones.

Notemos que como los parámetros iterativos $\varkappa_j$ satisfacen la estimación (véase el § 1, cap. XI) $\eta < \varkappa_j < 1$, entonces los parámetros $\omega_j^{(1)}$ y $\omega_j^{(2)}$ pertenecen al intervalo (δ, Δ). Por eso los operadores $\omega_k^{(\alpha)} E + A_\alpha$ son definidos positivos en H, y para su inversión se puede utilizar el algoritmo de factorización tripuntual usual.

2. **Método directo para el problema de Neumann.** Examinemos ahora el método directo —combinación del método de separación de variables y del método de reducción— para resolver el problema de diferencias (3). Recordemos que este método fué construido en el punto 2 del § 3 del cap. IV para el siguiente problema de contorno: en la región G se da la ecuación (1), sobre los lados $x_2 = 0$ y $x_2 = l_2$ están dadas las condiciones de contorno (2), y sobre los lados $x_1 = 0$ y $x_1 = l_1$ en lugar de las condiciones de segundo género (2) fueron dadas las condiciones de contorno de tercer género

$$\frac{\partial u}{\partial x_1} = \varkappa_{-1} u - g_{-1}(x_1), \quad x_1 = 0,$$

$$-\frac{\partial u}{\partial x_2} = \varkappa_{+1} u - g_{+1}(x_1), \quad x_1 = l_1,$$

al mismo tiempo $\varkappa_{-1}$ y $\varkappa_{+1}$ son las constantes no negativas y simultáneamente no iguales a cero. El correspondiente

problema de diferencias se diferencia del problema (3) solamente por la definición del operador Λ_1. Allí nosotros tuvimos ver con el operador Λ_1:

$$\Lambda_1 y = \begin{cases} \frac{2}{h_1}(y_{x_1} - \varkappa_{-1} y), & x_1 = 0, \\ y_{\bar{x}_1 x_1}, & h_1 \leqslant x_1 \leqslant l_1 - h_1, \\ \frac{2}{h_1}(-y_{\bar{x}_1} - \varkappa_{+1} y), & x_1 = l_1 \end{cases}$$

La exigencia de no anulación simultánea de $\varkappa_{-1}$ y $\varkappa_{+1}$ garantizó la solubilidad del problema de diferencias y la unicidad de la solución. En el algoritmo del método esta exigencia se utilizó sólo durante la resolución de los problemas de contorno tripuntuales para los coeficientes de Fourier de la solución buscada. Por eso para resolver el problema (3) se puede aprovechar formalmente el algoritmo citado en el punto 2 del § 3 del cap. IV, poniendo en él $\varkappa_{-1} = \varkappa_{+1} = 0$, y discutir la pregunta sobre la resolución de los problemas de contorno tripuntuales que aparecen.

Regresemos al problema (3). Vamos a considerar que $f \perp \ker A$, es decir, $(f, 1) = 0$. Entonces el problema es soluble, la solución normal $\overline{u}$ es ortogonal al ker A, y una de las posibles soluciones se puede separar, fijando su valor en un nodo de la red $\overline{\omega}$. En el algoritmo a examinar es cómodo realizar el desprendimiento de una de las posibles soluciones, fijando no la propia solución en un nodo, sino uno de los coeficientes de Fourier. Sea $y(i, j)$ una solución del problema (3). Entonces la solución normal $\overline{u}$ se puede hallar mediante la fórmula

$$\overline{u} = y - (y, \mu_{00})\,\mu_{00}, \quad \mu_{00}(i, j) = 1/\sqrt{l_1 l_2} \tag{7}$$

Mostremos ahora el algoritmo del método directo de resolución del problema de Neumann (3) para la ecuación de Poisson en un rectángulo.

1) Para $0 \leqslant i \leqslant N_1$ se calculan los valores de la función

$$\varphi(i, j) = \begin{cases} 2[f(i, 0) + f(i, 1)] - h_2^2 \Lambda_1 f(i, 0), & j = 0, \\ f(i, 2j-1) + f(i, 2j+1) + 2f(i, 2j) - \\ \qquad - h_2^2 \Lambda_1 f(i, 2j), & 1 \leqslant j \leqslant M_2 - 1, \\ 2[f(i, N_2) + f(i, N_2 - 1)] - h_2^2 \Lambda_1 f(i, N_2), \\ \qquad j = M_2, \quad \varkappa_{-1} = \varkappa_{+1} = 0. \end{cases}$$

2) Mediante el algoritmo de la transformación de Fourier rápida se calculan los coeficientes de Fourier de la función $\varphi(i, j)$:

$$z_{k_2}(i) = \sum_{j=0}^{M_2} \rho_j \varphi(i, j) \cos \frac{k_2 \pi j}{M_2}, \quad 0 \leqslant k_2 \leqslant M_2, \quad 0 \leqslant i \leqslant N_1.$$

3) Se resuelven los problemas de contorno tripuntuales

$$4 \operatorname{sen}^2 \frac{k_2 \pi}{2N_2} w_{k_2}(i) - h_2^2 \Lambda_1 w_{k_2}(i) = h_2^2 z_{k_2}(i), \quad 0 \leqslant i \leqslant N_1,$$

$$4 \cos^2 \frac{k_2 \pi}{2N_2} y_{k_2}(i) - h_2^2 \Lambda_1 y_{k_2}(i) = w_{k_2}(i), \quad 0 \leqslant i \leqslant N_1 \tag{8}$$

para $0 \leqslant k_2 \leqslant M_2$ y como resultado se encuentran los coeficientes de Fourier $y_{k_2}(i)$ de la función $y(i, j)$.

4) Por el algoritmo de la transformación de Fourier rápida se encuentra la solución del problema sobre las filas pares de la red $\bar{\omega}$

$$y(i, 2j) = \sum_{k_2=0}^{M_2} \rho_{k_2} y_{k_2}(i) \cos \frac{k_2 \pi j}{M_2},$$

$$0 \leqslant j \leqslant M_2, \quad 0 \leqslant i \leqslant N_1,$$

y se resuelven los problemas de contorno tripuntuales

$$2y(i, 2j-1) - h_2^2 \Lambda_1 y(i, 2j-1) =$$

$$= h_2^2 f(i, 2j-1) + u(i, 2j-2) + u(i, 2j),$$

$$0 \leqslant i \leqslant N_1, \quad 1 \leqslant j \leqslant M_2$$

para encontrar la solución sobre las filas impares.

Aquí son utilizadas las notaciones

$$M_2 = 0{,}5N_2, \quad \rho_j = \begin{cases} 1, & 1 \leqslant j \leqslant M_2 - 1; \\ 0{,}5, & j = 0, M_2, \end{cases}$$

el operador Λ_1 está definido en (4) y se supone que N_2 es una potencia de 2. El número de operaciones del método descrito será igual a $O(N^2 \log_2 N)$ para $N_1 = N_2 = N$.

La separación de una solución del conjunto de soluciones del problema (3) en el algoritmo citado se realiza de la manera siguiente. De todos los problemas de contorno tripuntuales que hay que resolver, únicamente un problema (8) para $k_2 = 0$ tiene solución no única. La separación aquí de una de las soluciones asegura la solución del problema

planteado. El problema de diferencias (8) para $k_2 = 0$ tiene la forma

$$\Lambda_1 w_0 (i) = -z_0 (i), \quad 0 \leqslant i \leqslant N_1,$$

ó

$$\begin{aligned} &(w_0)_{\bar{x}_1 x_1} = -z_0 (i), \quad 1 \leqslant i \leqslant N_1 - 1 \\ &\frac{2}{h_1} (w_0)_{x_1} = -z_0 (0), \quad i = 0, \\ &-\frac{2}{h_1} (w_0)_{\bar{x}_1} = -z_0 (N_1), \quad i = N_1. \end{aligned} \tag{9}$$

No es complejo mostrar, utilizando la ortogonalidad de $f(i, j)$ a $\mu_{00}(i, j)$, que la función reticular $z_0 (i)$ es ortogonal a la función $\mu_0 (i) = 1/\sqrt{l_1}$ en el sentido del producto escalar

$$(u, v)_1 = \sum_{x_1 = 0}^{l_1} u(x_1) v(x_1) h_1 (x_1).$$

Y como $\mu_0 (i)$ es la base en el subespacio ker Λ_1, entonces el problema (9) tiene solución. Separemos una de las soluciones, fijando el valor $w_0 (i)$ para algún i, $0 \leqslant i \leqslant N_1$. Pongamos, por ejemplo, $w_0 (N_1) = 0$ y eliminemos de (9) la condición de contorno para $i = N_1$. El problema de diferencias obtenido como resultado de esta sustitución se resuelve fácilmente por el método de factorización.

Después de que sea hallada una de las soluciones $y(i, j)$ del problema (3) según el algoritmo descrito más arriba, la solución normal $\bar{u}$, si se necesita, se determina por la fórmula (7).

Como conclusión señalemos que un procedimiento análogo de separación de una de las posibles soluciones puede ser utilizado también en el método de reducción completa, cuando se utiliza para resolver el problema de Neumann de diferencias.

3. Esquema iterativo con un operador degenerado B. La presencia de los métodos directos de inversión del operador de Laplace en un rectángulo en el caso de las condiciones de contorno de Neumann permite utilizar tales operadores en calidad del operador B en los esquemas iterativos implícitos de resolución de ecuaciones degeneradas. Ya que en este caso el operador B es degenerado, entonces es necesario estudiar de nuevo el problema de la elección de los parámetros iterativos.

Examinemos los métodos iterativos de resolución de la ecuación (5) bajo las siguientes suposiciones: 1) el operador A

es autoconjugado y degenerado; 2) es conocido el núcleo del operador A, es decir, está dada una base en el ker A; 3) el segundo miembro f de la ecuación (5) pertenece a im A, es decir, $f = \overline{f} \in \text{im } A$. Esta condición es fácil de satisfacer ya que está conocida la base en el ker A. Con esto la solución normal $\overline{u}$ de la ecuación (5) es clásica y satisface la relación

$$A\overline{u} = f. \tag{10}$$

Señalemos que en virtud de la autoconjugación del operador A tiene lugar la siguiente descomposición ortogonal del espacio H:

$$H = \ker A \oplus \text{im } A. \tag{11}$$

Para resolver la ecuación (5) examinaremos el esquema implícito de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, y_0 \in H, \tag{12}$$

con el operador degenerado B. Se plantea el problema de hallar con ayuda de (12) la aproximación a una de las soluciones de la ecuación (5).

Formularemos ahora las suposiciones complementarias con respecto a los operadores A y B. Sea B un operador autoconjugado en H y ker B = ker A. Además, supongamos que para cualquier $x \in \text{im } A$ se cumplen las desigualdades

$$\gamma_1 (Bx, x) \leqslant (Ax, x) \leqslant \gamma_2 (Bx, x), \quad \gamma_1 > 0, \tag{13}$$
$$Ax \neq 0, \quad (Bx, x) > 0.$$

Notemos que de las condiciones $B = B^*$, ker B = ker A y (11) se deduce la coincidencia de im A e im B.

Estudiemos el esquema (12). En correspondencia con (11) representaremos y_k en forma de la suma

$$y_k = \overline{y}_k + \widetilde{y}_k, \quad \overline{y}_k \in \text{im } A, \quad \widetilde{y}_k \in \ker A.$$

De (12) obtenemos la siguiente ecuación para y_{k+1}:

$$By_{k+1} = \varphi_k \tag{14}$$

donde $\varphi_k = By_k - \tau_{k+1}(Ay_k - f)$.

Puesto que $f \in \text{im } A$ e im B = im A, entonces $\varphi_k \in$ $\in$ im A para cualquier y_k. Por consiguiente $\varphi_k \perp \ker B$, y la ecuación (14) tiene un conjunto de soluciones en el sentido usual, y su solución normal $\overline{y}_{k+1}$ satisface la ecua-

ción

$$B\bar{y}_{k+1} - \varphi_k. \tag{15}$$

Subrayamos que en virtud de las igualdades $B\tilde{y}_k = A\tilde{y}_k = 0$ tenemos

$$\varphi_k = B\bar{y}_k - \tau_{k+1}(A\bar{y}_k - f). \tag{16}$$

Por lo tanto la componente $\tilde{y}_k$ de la aproximación iterativa y_k, $\tilde{y}_k \in \ker A$, no ejerce ninguna influencia sobre $\bar{y}_{k+1}$. De aquí se desprende una conclusión: al resolver la ecuación (14) es suficiente hallar alguna solución suya y sólo después de finalizar el proceso de iteraciones calcular la proyección de y_n sobre $\operatorname{im} A$, es decir hallar $\bar{y}_n$.

Examinemos ahora la pregunta sobre la elección del parámetro iterativo τ_k. En virtud de lo dicho más arriba se debe elegirlo de manera tal, que la sucesión $\bar{y}_k$ tiende a la solución normal $\bar{u}$ de la ecuación (5). De (10), (15) y (16) obtenemos el siguiente problema para el error $\bar{z}_k = \bar{y}_k - \bar{u}$:

$$B\bar{z}_{k+1} = (B - \tau_{k+1}A)\bar{z}_k, \quad k = 0, 1, \ldots, \tag{17}$$

donde $\bar{z}_k \in \operatorname{im} A$ para cualquier $k \geqslant 0$.

Como en el subespacio $\operatorname{im} A$ los operadores A y B en virtud de (13) son definidos positivos, entonces el esquema se le puede investigar, en la forma usual, de convergencia según la norma del espacio energético H_D, donde $D = A$, B ó $AB^{-1}A$. Puesto que en este caso el operador $DB^{-1}A$ es autoconjugado, los parámetros τ_k se pueden elegir por las fórmulas del método iterativo de Chebishev (véase el § 2 del cap. VI).

$$\tau_k = \frac{\tau_0}{1 + \rho_0 \mu_k}, \quad \mu_k \in \mathfrak{M}_n^* = \left\{\cos\left(\frac{(2i-1)\pi}{2n}, \quad 1 \leqslant i \leqslant n\right\}\right.,$$

$$1 \leqslant k \leqslant n,$$

$$\tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \rho_1 = \frac{1-\sqrt{\varepsilon}}{1+\sqrt{\varepsilon}}, \quad \xi = \frac{\gamma_1}{\gamma_2}, \tag{18}$$

$$n \geqslant n_0(\varepsilon) = \ln(0{,}5\varepsilon)/\ln \rho_1,$$

utilizando γ_1 y γ_2 de las desigualdades (13). Entonces para el error $\bar{z}_n$ después de n iteraciones será cierta la estimación

$$\|\bar{z}_n\|_D \leqslant \varepsilon \|\bar{z}_0\|_D.$$

El sentido del examen de los métodos iterativos con un operador degenerado B consiste en lo siguiente. Si el operador B es tal, que la obtención de la solución de la ecuación (14) se realiza de forma considerablemente más simple que la de la ecuación (5), y si la relación ξ no es demasiado pequeña, entonces este método de resolución aproximada de la ecuación (5) pueda aparecerse útil.

Citemos el ejemplo de un problema de diferencias, sobre el cual ilustraremos el método propuesto. Supongamos que sobre la red rectangular

$$\overline{\omega} = \{x_{ij} = (ih_1, jh_2) \in \overline{G}, \quad 0 \leqslant i \leqslant N_1, \quad 0 \leqslant j \leqslant N_2, \quad h_\alpha N_\alpha = l_\alpha, \quad \alpha = 1, 2\},$$

introducida en el rectángulo $\overline{G}$, se exige hallar la solución del problema de Neumann para una ecuación elíptica con coeficientes variables

$$\Lambda y = -f(x), \quad x \in \overline{\omega} \tag{19}$$

$$\Lambda = \Lambda_1 + \Lambda_2, \quad f(x) = \varphi(x) + \frac{2}{h_1}\varphi_1(x) + \frac{2}{h_2}\varphi_2(x),$$

donde

$$\Lambda_\alpha y = \begin{cases} \frac{2}{h_\alpha} a_\alpha^{+1} y_{x_\alpha}, & x_\alpha = 0, \\ (a_\alpha y_{\overline{x}_\alpha})_{x_\alpha}, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ -\frac{2}{h_\alpha} a_\alpha y_{\overline{x}_\alpha}, & x_\alpha = l_\alpha, \end{cases}$$

$$\varphi_\alpha(x_\alpha) = \begin{cases} g_{-\alpha}(x_\beta), & x_\alpha = 0, \\ 0, & h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha, \\ g_{+\alpha}(x_\beta), & x_\alpha = l_\alpha, \end{cases}$$

$$a_1^{+1}(x) = a_1(x_1 + h_1, x_2), \quad a_2^{+1}(x) = a_2(x_1, x_2 + h_2).$$

Se supone que los coeficientes $a_1(x)$ y $a_2(x)$ satisfacen las condiciones

$$0 < c_1 \leqslant a_\alpha(x) \leqslant c_2, \quad \alpha = 1, 2, \quad x \in \omega. \tag{20}$$

El esquema (19) es el análogo de diferencias del problema de Neumann para la ecuación elíptica.

$$\frac{\partial}{\partial x_1}\left(k_1(x)\frac{\partial u}{\partial x_1}\right)+\frac{\partial}{\partial x_2}\left(k_2(x)\frac{\partial u}{\partial x_2}\right)=-\varphi(x), \quad x\in G,$$

$$k_\alpha \frac{\partial u}{\partial x_\alpha}=-g_{-\alpha}(x_\beta), \quad x_\alpha=0, \quad \beta=3-\alpha,$$

$$-k_\alpha \frac{\partial u}{\partial x_\alpha}=-g_{+\alpha}(x_\beta), \quad x_\alpha=l_\alpha, \quad \alpha=1, 2.$$

El espacio H está definido en el punto 1. Introduciendo el operador $A=-\Lambda$, escribamos el problema de diferencias (19) en forma de la ecuación (5). Es fácil comprobar que $A=A^*$, y las fórmulas de Green de diferencias dan

$$(Ay, y)=\sum_{\alpha=1}^{2}(a_\alpha y_{\bar{x}_\alpha}^2, 1)_\alpha, \tag{21}$$

donde son utilizadas las siguientes notaciones:

$$(u, v)_\alpha=\sum_{x_\beta=0}^{l_\beta}\sum_{x_\alpha=h_\alpha}^{l_\alpha}u(x)v(x)\hbar_\beta(x_\beta)h_\alpha,$$

$$\beta=3-\alpha, \quad \alpha=1, 2.$$

No es difícil mostrar, que el operador A es degenerado y para cualesquiera coeficientes $a_\alpha(x)$ que satisfagan (20), el núcleo del operador A lo componen las funciones reticulares que son constantes sobre $\bar{\omega}$. Por eso en calidad de base en el ker A puede ser tomada la función conocida por nosotros $\mu_{00}(i, j)=1/\sqrt{l_1 l_2}$.

Definamos ahora el operador $B=-\mathring{\Lambda}$, donde $\mathring{\Lambda}=\mathring{\Lambda}_1+\mathring{\Lambda}_2$,

$$\mathring{\Lambda}_\alpha y=\begin{cases}\frac{2}{h_\alpha}y_{x_\alpha}, & x_\alpha=0,\\ y_{\bar{x}_\alpha x_\alpha}, & h_\alpha\leqslant x_\alpha\leqslant l_\alpha-h_\alpha,\\ -\frac{2}{h_\alpha}y_{\bar{x}_\alpha}, & x_\alpha=l_\alpha, \quad \alpha=1, 2.\end{cases}$$

El operador B es autoconjugado en H, y en el punto 1 fue señalado, que precisamente la función $\mu_{00}(i, j)$ forma una base en el ker B. Por consiguiente, ker A es conocido, y, por lo tanto, ker $A=$ ker B. Si además tomamos la proyección de f sobre im A y sustituimos, en caso de necesidad, el se-

gundo miembro en el esquema (19), todas las exigencias presentadas al esquema (12) y a la ecuación (5) serán cumplidas.

Para la aplicación del método iterativo (12), (18) queda indicar γ_1 y γ_2 en las desigualdades (13). Ya que

$$(By, y) = \sum_{\alpha=1}^{2} \left(y^2_{\bar{x}_\alpha}, 1\right)_\alpha, \tag{22}$$

y los subespacios im A e im B coinciden y consisten de las funciones reticulares que no son constantes sobre la red $\bar{\omega}$, entonces de (20)—(22) obtenemos que $\gamma_1 = c_1$ y $\gamma_2 = c_2$. La información a priori necesaria está hallada.

De la estimación (18) para el número de iteraciones se ve que la misma no depende del número de incógnitas en el problema, y está determinado solamente por la relación c_1/c_2. Luego en virtud de la elección del operador B la ecuación (14) para y_{k+1} es el problema de Neumann de diferencias para la ecuación de Poisson en un rectángulo. Su solución se puede hallar por el método directo expuesto en el punto 2 con un gasto de $O(N^2 \log_2 N)$ operaciones aritméticas. Entonces el número total de operaciones que conviene gastar en el método propuesto para obtener la solución (19) con exactitud ε, será igual a $Q(\varepsilon) = O(N^2 \log_2 N \mid \ln \varepsilon \mid)$.

Capítulo XIII

Métodos iterativos de resolución de ecuaciones no lineales

En este capítulo se estudian los métodos iterativos de resolución de esquemas de diferencias no lineales. En el § 1 se expone la teoría general de los métodos iterativos para una ecuación operacional no lineal abstracta en un espacio de Hilbert siendo diferentes suposiciones respecto al operador. En el § 2 se examina la aplicación de la teoría general a la resolución de los análogos de diferencias de los problemas de contorno para las ecuaciones elípticas cuasilineales de segundo orden.

§ 1. Métodos iterativos. La teoría general

1. Método de iteración simple para las ecuaciones con un operador monótono. En los capítulos anteriores fueron estudiados los métodos iterativos de resolución de la ecuación lineal operacional de primer género

$$Au = f, \tag{1}$$

dada en un espacio de Hilbert H. La mayoría de los métodos construidos eran lineales y convergían con la velocidad de progresión geométrica.

Pasemos ahora al estudio de los métodos de resolución de la ecuación (1) en el caso, cuando A es un operador no lineal arbitrario, el cual actúa en H. Este capítulo está dedicado a la construcción de métodos iterativos para resolver las ecuaciones no lineales (1). La construcción de tales métodos se basa, por lo general, en el empleo del operador lineal B en los esquemas iterativos implícitos, el cual en cierto sentido es cercano al operador no lineal A. Más abajo para diferentes suposiciones respecto a los operadores A, B y D serán demostrados los teoremas generales de convergencia

en H_D de la solución del esquema iterativo implícito de dos capas

$$B\frac{y_{k+1}-y_k}{\tau}+Ay_k=f, \quad k=0, 1, \ldots, y_0\in H. \tag{2}$$

Comenzaremos el estudio del esquema iterativo (2) por el caso de un operador A monótono. Recordemos que un operador A, definido en un espacio de Hilbert real, se llama *monótono*, si

$$(Au-Av, u-v)\geqslant 0, \quad u, v\in H,$$

y *fuertemente monótono*, si existe un número $\delta>0$ tal, que para cualesquiera $u, v\in H$

$$(Au-Av, u-v)\geqslant \delta \|u-v\|^2. \tag{3}$$

Del teorema 11 del capítulo V se sigue la existencia y la unicidad de solución, en la bola $\|u\|\leqslant \frac{1}{\delta}\|A0-f\|$, de la ecuación (1) de un operador fuertemente monótono y continuo en el espacio de dimensión finita H.

Supondremos que B es el operador lineal acotado y definido positivo en H, y D es el operador autoconjugado y definido positivo en H. Sean, además, prefijadas las constantes γ_1 y γ_2 en las desigualdades

$$(DB^{-1}(Au-Av), B^{-1}(Au-Av))\leqslant \gamma_2 (DB^{-1}(Au-Av), u-v), \tag{4}$$

$$(DB^{-1}(Au-Av), u-v)\geqslant \gamma_1 (D(u-v), u-v), \tag{5}$$

y al mismo tiempo $\gamma_1>0$.

LEMA 1. *Supongamos que se cumplen las condiciones* (4) *y* (5). *Entonces la ecuación* (1) *es resoluble unívocamente para todo segundo miembro.*

En efecto, escribamos la ecuación (1) en la forma equivalente

$$u=Su, \tag{6}$$

donde el operador no lineal S se define de la manera siguiente:

$$Su=u-\tau B^{-1}Au+\tau B^{-1}f, \quad \tau>0.$$

Mostremos que en H_D para $\tau\leqslant 2/\gamma_2$ el operador S es uniformemente contractante, es decir, para cualesquiera $u, v\in H$

es válida la estimación

$$\| Su - Sv \|_D \leqslant \rho(\tau) \| u - v \|_D, \quad \rho(\tau) \leqslant 1, \qquad (7)$$

además $\rho(\tau)$ no depende de u ni de v. Entonces la afirmación del lema se deducirá del teorema 8 del capítulo V sobre las aplicaciones contraídas.

Tenemos

$$\| Su - Sv \|_D^2 = (D(Su - Sv), \quad (Su - Sv)) =$$
$$= \| u - v \|_D^2 - 2\tau (DB^{-1}(Au - Av), u - v) +$$
$$+ \tau^2 (DB^{-1}(Au - Av), B^{-1}(Au - Av)).$$

De (4) y (5) hallaremos para $\tau < 2/\gamma_2$,

$$\|Su - Sv\|_D^2 \leqslant \|u - v\|_D^2 - \tau(2 - \tau\gamma_2)(DB^{-1}(Au - Av),$$
$$u - v) \leqslant \rho^2(\tau) \|u - v\|_D^2,$$

donde

$$\rho^2(\tau) = 1 - \tau(2 - \tau\gamma_2)\gamma_1. \qquad (8)$$

Puesto que $\tau < 2/\gamma_2$, entonces $\rho(\tau) < 1$. El lema está demostrado.

Investiguemos ahora la convergencia del esquema iterativo (2) bajo la suposición de que están cumplidas las condiciones (4) y (5).

De (2) hallaremos

$$y_{k+1} = y_k - \tau B^{-1} A y_k + \tau B^{-1} f = S y_k, \qquad (9)$$

donde el operador no lineal S está definido más arriba. Ya que la solución u de la ecuación (1) satisface la relación (6), entonces de (6)—(9) obtenemos

$$y_{k+1} - u = Sy_k - Su, \quad k = 0, 1, \ldots,$$
$$\|y_{k+1} - u\|_D^2 = \|Sy_k - Su\|_D^2 \leqslant \rho^2(\tau) \|y_k - u\|_D^2,$$

donde $\rho^2(\tau)$ está definido en (8). No es difícil ver que la mejor estimación de la velocidad de convergencia se alcanza, cuando $\rho(\tau)$ es mínimo, o sea, para $\tau = \tau_0 = 1/\gamma_2$. Con esto $\rho_0 = \rho(\tau_0) = \sqrt{1 - \xi}$, $\xi = \gamma_1/\gamma_2$. Así, está demostrado

TEOREMA 1. *Sean cumplidas las condiciones* (4) *y* (5). *El método iterativo* (2) *con* $\tau = \tau_0 = 1/\gamma_2$ *converge en* H_D, *y para el error tiene lugar la estimación*

$$\|y_n - u\|_D \leqslant \rho_0^n \|y_0 - u\|_D, \quad \rho_0 = \sqrt{1 - \xi}, \quad \xi = \gamma_1/\gamma_2,$$

donde u es la solución de la ecuación (1). *Para el número de iteraciones es cierta la estimación*

$$n \geqslant n_0(\varepsilon) = \ln \varepsilon / \ln \rho_0.$$

Notemos que para un operador lineal A las condiciones (4) y (5) se pueden escribir en la forma

$$(DB^{-1}Ay, B^{-1}Ay) \leqslant \gamma_2 (DB^{-1}Ay, y), \quad (DB^{-1}Ay, y) \geqslant \geqslant \gamma_1 (Dy, y).$$

Por consiguiente, en este caso ellos coinciden con las condiciones impuestas a los operadores A, B y D, cuando el operador $DB^{-1}A$ es no autoconjugado en H. Con esto el método aquí construido se convierte en la primera variante del método de iteración simple para el caso no autoconjugado (véase el punto 2 del § 4 del cap. VI).

Señalemos que en lugar de (4) se puede exigir el cumplimiento de la condición

$$\|B^{-1}(Au - Av)\|_D \leqslant \bar{\gamma}_2 \|u - v\|_D, \tag{10}$$

la cual para $D = B = E$ es la condición de Lipzchitz para el operador A. De (10) y (5) se sigue la desigualdad (4) con $\gamma_2 = \bar{\gamma}_2^2/\gamma_1$.

Si el operador B es autoconjugado y definido positivo en H, entonces en calidad de operador D es posible tomar B. Entonces las condiciones (4) y (5) tendrán la forma

$$(B^{-1}(Au - Av), Au - Av) \leqslant \gamma_2 ((Au - Av), u - v),$$
$$(Au - Av, u - v) \geqslant \gamma_1 (B(u - v), u - v), \quad \gamma_1 > 0.$$

Si B es no autoconjugado y no degenerado, entonces para $D = B^*B$ las condiciones (4) y (5) tienen la forma

$$(Au - Av, Au - Av) \leqslant \gamma_2 (Au - Av, B(u - v)),$$
$$(Au - Av, B(u - v)) \geqslant \gamma_1 (B(u - v), B(u - v)),$$
$$\gamma_1 > 0.$$

Para $D = B = E$ la condición (5) significa, que el operador A debe ser fuertemente monótono en H.

2. Métodos iterativos para el caso del operador diferenciable. Un mejoramiento de la estimación de la velocidad de convergencia del método de iteración simple para la ecuación (1) puede ser alcanzado a cuenta de más fuertes restricciones sobre el operador A. Precisamente, vamos a considerar que el operador A tiene derivada de Gato. Re-

cordemos que el operador lineal $A'(u)$ se llama *derivada de Gato* del operador A en el punto $u \in H$, si para cualquier $v \in H$ es válida la relación

$$\lim_{t \to 0} \left\| \frac{A(u+tv) - A(u)}{t} - A'(u)\, v \right\| = 0.$$

Si el operador A tiene derivada de Gato en cada punto del espacio H, entonces es válida la desigualdad de Lagrange

$$\|Au - Av\| \leqslant \sup_{0 \leqslant t \leqslant 1} \|A'(u + t(v-u))\|\, \|u - v\|, \quad u,\ v \in H,$$

y para cualesquiera u, v y $w \in H$ existe $t \in [0, 1]$ tal que

$$(Au - Av,\ w) = (A'(u + t(v-u))\, z,\ w), \quad z = u - v. \quad (11)$$

Regresemos a la investigación de la convergencia del esquema iterativo (2). Tiene lugar

TEOREMA 2. *Supongamos que el operador A tiene derivada de Gato $A'(v)$ en la bola $\Omega(r) = \{v : \| u - v \|_D \leqslant r\}$, la cual para cualquier $v \in \Omega(r)$ satisface las desigualdades*

$$(DB^{-1}A'(v)\, y,\ B^{-1}A'(v)\, y) \leqslant \gamma_2 (DB^{-1}A'(v)\, y,\ y), \quad (12)$$

$$(DB^{-1}A'(v)\, y,\ y) \geqslant \gamma_1 (Dy,\ y), \quad \gamma_1 > 0$$

para cualquier $y \in H$. El método iterativo (2) con $\tau = 1/\gamma_2$ e $y_0 \in \Omega(r)$ converge en H_D, y para el error es cierta la estimación

$$\| y_n - u \|_D \leqslant \rho^n \| y_0 - u \|_D, \quad (13)$$

donde u es la solución de la ecuación (1), $\rho = \sqrt{1 - \xi}$, $\xi = \gamma_1/\gamma_2$. Si el operador $DB^{-1}A'(v)$ es autoconjugado en H para $v \in \Omega(r)$ y si se cumplen las desigualdades

$$\gamma_1 (Dy,\ y) \leqslant (DB^{-1}A'(v)\, y,\ y) \leqslant \gamma_2 (Dy,\ y), \quad \gamma_1 > 0 \quad (14)$$

para cualquier $v \in \Omega(r)$ e $y \in H$, entonces siendo $\tau = \tau_0 = 2/(\gamma_1 + \gamma_2)$ para el proceso iterativo (2) es cierta la estimación (13) si $\rho = \rho_0 = (1 - \xi)/(1 + \xi)$.

En efecto, de la ecuación para el error

$$y_{k+1} - u = Sy_k - Su, \quad Sv = v - \tau B^{-1}Av + \tau B^{-1}f$$

y de la desigualdad de Lagrange obtenemos

$$\| y_{k+1} - u \|_D = \| Sy_k - Su \|_D \leqslant$$
$$\leqslant \sup_{0 \leqslant t \leqslant 1} \| S'(v_k) \|_D \| y_k - u \|_D, \quad (15)$$

donde $v_k = y_k + t(u - y_k) \in \Omega(r)$, si $y_k \in \Omega(r)$. Ya que $S'(v_k) = E - \tau B^{-1}A'(v_k)$, entonces el problema se reduce a la estimación en H_D de la norma del operador lineal $E - \tau B^{-1}A'(v_k)$. De la definición de la norma del operador tenemos

$$\|S'(v_k)\|_D^2 = \sup_{y \neq 0} \frac{(S'(v_k)\, y,\ S'(v_k)\, y)_D}{(y,\ y)_D} =$$
$$= \sup_{y \neq 0} \frac{(DS'(v_k)\, y,\ S'(v_k)\, y)}{(Dy,\ y)} = \sup_{z \neq 0} \times$$
$$\times \frac{((E - \tau C(v_k))\, z,\ (E - \tau C(v_k))\, z)}{(z,\ z)} = \|E - \tau C(v_k)\|^2,$$

donde $C(v_k) = D^{-1/2}(DB^{-1}A'(v_k))\, D^{-1/2}$ y fue cumplida la sustitución $y = D^{-1/2}z$.

Sustituyendo la relación encontrada en (15), obtenemos

$$\|y_{k+1} - u\|_D \leqslant \sup_{0 \leqslant t \leqslant 1} \|E - \tau C(v_k)\| \, \|y_k - u\|_D.$$

De (12) hallaremos que el operador $C(v_k)$ para toda $v_k \in \Omega(r)$ satisface las desigualdades

$$(C(v_k)\, y,\ C(v_k)\, y) \leqslant \gamma_2 (C(v_k)\, y,\ y),$$
$$(C(v_k)\, y,\ y) \geqslant \gamma_1 (y,\ y).$$

Recordemos que la estimación requerida para la norma del operador lineal $E - \tau C(v_k)$ bajo las suposiciones indicadas fue obtenida en el punto 2 del § 4 del cap. VI. Precisamente, para $\tau = 1/\gamma_2$ tenemos $\| E - \tau C(v_k) \| \leqslant \rho$, donde $\rho = \sqrt{1 - \xi}$, $\xi = \gamma_1/\gamma_2$. La primera afirmación del teorema está demostrada. De modo análogo se demuestra también la segunda. En este caso el operador $C(v_k)$ es autoconjugado en H, y la estimación para la norma del operador $E - \tau C(v_k)$ fue obtenida anteriormente en el punto 2 del § 3 del cap. VI. El teorema 2 está demostrado.

En el capítulo VI, además de la estimación aquí utilizada para la norma del operador $E - \tau C(v_k)$ en el caso no autoconjugado, fue obtenida otra estimación partiendo de la suposición de que sean dados los tres números $\bar{\gamma}_1$, $\bar{\gamma}_2$ y $\bar{\gamma}_3$

en las desigualdades

$$\bar{\gamma}_1 E \leqslant C(v_h) \leqslant \bar{\gamma}_2 E, \quad \|C_1(v_h)\| \leqslant \bar{\gamma}_3, \quad \bar{\gamma}_1 > 0,$$

donde $C_1 = 0{,}5\,(C - C^*)$ es la parte antisimétrica del operador C. En este caso para $\tau = \tau_0(1 - \varkappa\bar{\rho})$ es válida la estimación $\|E - \tau C(v_h)\| \leqslant \bar{\rho}$, donde

$$\varkappa = \frac{\bar{\gamma}_3}{\sqrt{\bar{\gamma}_1\bar{\gamma}_2 + \bar{\gamma}_3^2}}, \quad \tau_0 = \frac{2}{\bar{\gamma}_1 + \bar{\gamma}_2}, \quad \bar{\rho} = \frac{1-\bar{\xi}}{1+\bar{\xi}}, \bar{\xi} = \frac{1-\varkappa}{1+\varkappa}\frac{\bar{\gamma}_1}{\bar{\gamma}_2}. \quad (16)$$

TEOREMA 3. *Supongamos que en la esfera* $\Omega(r)$ *el operador* A *posee derivada de Gato* $A'(v)$, *la cual para cualquier* $v \in \Omega(r)$ *satisface las desigualdades*

$$\bar{\gamma}_1(Dy, y) \leqslant (DB^{-1}A'(v)y, y) \leqslant \bar{\gamma}_2(Dy, y), \gamma_1 > 0,$$
$$\|0,\ 5\,(DB^{-1}A'(v) - A'^*(v)(B^*)^{-1}D)y\|^2_{D^{-1}} \leqslant \bar{\gamma}_3^2(Dy, y). \quad (17)$$

Entonces para $\tau = \tau_0(1 - \varkappa\bar{\rho})$ *e* $y_0 \in \Omega(r)$ *el método iterativo* (2) *converge en* H_D, *y para el error es cierta la estimación* (13), *donde* $\rho = \bar{\rho}$ *está definido en* (16).

Mostremos ahora que si el operador $A'(w)$ para $w \in \Omega(r)$ satisface las condiciones (17), entonces para cualesquiera $u, v \in \Omega(r)$ se cumplen las desigualdades (4) y (5) con las constantes $\gamma_1 = \bar{\gamma}_1$, $\gamma_2 = (\bar{\gamma}_2 + \bar{\gamma}_3)^2/\bar{\gamma}_1$. Entonces del lema 1 se deducirá la unicidad de la solubilidad de la ecuación (1).

En virtud de (11) tenemos para $u, v \in \Omega(r)$ y $t \in [0\ 1]$

$$(DB^{-1}Au - DB^{-1}Av, u - v) = (Ry, y), \quad R = DB^{-1}A'(w)$$

donde $y = u - v$, $w = u + t(v - u) \in \Omega(r)$. De (17) obtenemos $(Ry, y) \geqslant \bar{\gamma}_1(Dy, y)$, es decir, está cumplida la desigualdad (5) de $\gamma_1 = \bar{\gamma}_1$.

Luego tenemos que $(DB^{-1}Au - DB^{-1}Av, z) = (Ry, z)$. Representemos el operador R en forma de la suma $R = R_0 + R_1$, donde $R_0 = 0{,}5\,(R + R^*)$ y $R_1 = 0{,}5\,(R - R^*) = 0{,}5\,(DB^{-1}A'(w) - A'^*(w)(B^*)^{-1}D)$ son respectivamente la parte simétrica y la antisimétrica del operador R.

En virtud de la desigualdad de Cauchy — Buniakovski y de la condición (17) obtendremos

$$(R_1y, z) = (D^{-1/2}R_1y, D^{1/2}z) \leqslant (D^{-1}R_1y, R_1y)^{1/2}(Dz, z)^{1/2} =$$
$$= \|R_1y\|_{D^{-1}}(Dz, z)^{1/2} \leqslant \bar{\gamma}_3(Dy, y)^{1/2}(Dz, z)^{1/2}.$$

De la desigualdad generalizada de Cauchy—Buniakovski hallaremos

$$(R_0 y,\ z) \leqslant (R_0 y,\ y)^{1/2} (R_0 z,\ z)^{1/2} =$$
$$= (Ry,\ y)^{1/2} (Rz,\ z)^{1/2} \leqslant \bar{\gamma}_2 (Dy,\ y)^{1/2} (Dz,\ z)^{1/2}.$$

Así pues, nosotros obtenemos la desigualdad

$$(Ry,\ z) \leqslant (\bar{\gamma}_2 - \bar{\gamma}_3) (Dy,\ y)^{1/2} (Dz,\ z)^{1/2}.$$

Poniendo $z = B^{-1} (Au - Av)$ y utilizando (5), tendremos

$$(DB^{-1} (Au - Av),\ B^{-1} (Au - Av) \leqslant$$
$$\leqslant \frac{(\bar{\gamma}_2 + \bar{\gamma}_3)^2}{\bar{\gamma}_1} (DB^{-1} (Au - Av),\ u - v).$$

La afirmación está demostrada.

3. Método de Newton—Kantorovich. En los teoremas 2 y 3 nosotros supusimos que la derivada de Gato $A'(v)$ existe y satisface las desigualdades respectivas para $v \in \Omega(r) = \{v: \| u - v \|_D \leqslant r\}$ donde u es la solución de la ecuación (1).

De la demostración de teoremas se sigue que es suficiente exigir para cada iteración $k = 0, 1, \ldots$ el cumplimiento de dichas desigualdades sólo para $v \in \Omega(r_k)$, donde $r_k = \| u - y_k \|_D$.

En este caso γ_1 y γ_2 (y también $\bar{\gamma}_1$, $\bar{\gamma}_2$ y $\bar{\gamma}_3$) pueden depender del número de iteraciones k. Si elegimos el parámetro iterativo τ por las fórmulas de los teoremas 2 y 3, entonces obtendremos el proceso iterativo no estacionario (2) con $\tau = \tau_{k+1}$.

Además, se puede examinar el proceso iterativo

$$B_{k+1} \frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f,\ k = 0,\ 1,\ \ldots,\ y_0 \in H, \qquad (18)$$

en el cual el operador $B = B_{k+1}$ también depende del número de iteraciones. ¿Cómo elegir los operadores B_k? Si el operador A es lineal, entonces $A'(v) = A$ para cualquier $v \in H$. Entonces de los teoremas 2 y 3 se desprende que para $B = A'(v) = A$ la velocidad de convergencia del método iterativo (2) es máxima. Precisamente, para toda aproximación inicial y_0 obtendremos que $y_1 = u$.

Elijamos ahora el operador B_{k+1} en el caso de un operador A no lineal de la siguiente manera: $B_{k+1} = A'(y_k)$.

Obtendremos el esquema iterativo

$$A'(y_k)\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f,\ k=0,\ 1,\ \ldots,\ y_0\in H. \qquad (19)$$

En correspondencia con la terminología aceptada el proceso iterativo (19) es no lineal. Para $\tau_k\equiv 1$ él se llama *método de Newton—Kantorovich*. Para estimar la velocidad de convergencia del proceso iterativo (19) se pueden emplear los teoremas 2 y 3, en los cuales B conviene cambiar por $A'(y_k)$. En particular, para el caso $D=E$ con $\tau_{k+1}=1/\gamma_2$ tiene lugar la estimación

$$\|y_{k+1}-u\|\leqslant\rho\|y_k-u\|,\ \rho=\sqrt{1-\gamma_1/\gamma_2}<1, \qquad (20)$$

donde γ_1 y γ_2 son tomados de las desigualdades (12) del teorema 2

$$\|(A'(y_k))^{-1}A'(v)\,y\|^2\leqslant\gamma_2((A'(y_k))^{-1}A'(v)\,y,\ y)$$
$$((A'(y_k))^{-1}A'(v)\,y,\ y)\geqslant\gamma_1(y,\ y),\ \gamma_1>0$$

para $y\in H$, $v\in\Omega(r_k)$ y $r_k=\|u-y_k\|$. De (20) se deduce que $r_{k+1}=\|y_{k+1}-u\|\leqslant\rho r_k<r_k$ y, por consiguiente, $r_k\to 0$ cuando $k\to\infty$. Por eso, si la derivada $A'(v)$ como función de v de valores en el espacio de los operadores lineales es continua en el entorno de la solución, entonces cuando $k\to\infty$, tenemos que $\gamma_1\to 1$ y $\gamma_2\to 1$. Esto conduce a la aceleración de convergencia del método iterativo (19) al aumentar el número de la iteración k.

Los razonamientos citados muestran que los métodos del tipo (19) bajo ciertas suposiciones sobre la suavidad del operador $A'(v)$ poseen una velocidad de convergencia más alta que la velocidad de convergencia de una progresión geométrica.

Examinemos el método de Newton—Kantorovich (19) con $\tau_k\equiv 1$. Investigaremos la convergencia de este método bajo las siguientes suposiciones:

1) se cumplen las desigualdades

$$\|A'(v)-A'(w)\|\leqslant\alpha\|v-w\|,\quad \alpha\geqslant 0, \qquad (21)$$

$$\|A'(v)\,y\|\geqslant\frac{1}{\beta}\|y\|,\ y\in H,\ \beta>0 \qquad (22)$$

para $v,\ w\in\Omega(r)$; 2) la aproximación inicial y_0 pertenece a la esfera $\Omega(\bar{r})$, donde $\bar{r}=\min(r,\ 1/(\alpha\beta))$.

TEOREMA 4. *Si se cumplen las suposiciones* 1) *y* 2), *entonces para el error del método iterativo* (19) *con* $\tau_k\equiv 1$ *es*

cierta la estimación

$$\|y_n - u\| \leqslant \frac{1}{\alpha\beta}(\alpha\beta\|y_0 - u\|)^{2^n}. \qquad (23)$$

En efecto, de (19) obtenemos la relación siguiente:

$$A'(y_k)(y_{k+1} - u) = A'(y_k)(y_k - u) - (Ay_k - Au) =$$
$$= Ty_k - Tu,$$
$$Tu = A'(y_k)u - Au,$$

donde u es la solución de la ecuación (1). De aquí en virtud de la desigualdad de Lagrange para el operador lineal T obtendremos

$$\|A'(y_k)(y_{k+1} - u)\| = \|Ty_k - Tu\| \leqslant$$
$$\leqslant \sup_{0 \leqslant t \leqslant 1} \|T'(v_k)\| \, \|y_k - u\|,$$

donde $v_k = y_k + t(u - y_k)$. De la definición del operador T' tendremos

$$T'(v_k) = A'(y_k) - A'(v_k).$$

Supongamos que $y_k \in \Omega(\bar{r})$. Puesto que $\bar{r} \leqslant r$, entonces $y_k \in \Omega(r)$, y, por consiguiente, $v_k \in \Omega(r)$. De la desigualdad (21) tendremos

$$\|T'(v_k)\| = \|A'(y_k) - A'(v_k)\| \leqslant \alpha \|y_k - v_k\| =$$
$$= \alpha t \|y_k - u\|,$$
$$\sup_{0 \leqslant t \leqslant 1} \|T'(v_k)\| \leqslant \alpha \|y_k - u\|.$$

De esta forma, está hallada la estimación

$$\|A'(y_k)(y_{k+1} - u)\| \leqslant \alpha \|y_k - u\|^2.$$

Utilizando la desigualdad (22), de aquí obtendremos

$$\|y_{k+1} - u\| \leqslant \alpha\beta \|y_k - u\|^2. \qquad (24)$$

Notemos que como $\|y_k - u\| \leqslant \bar{r}$ y $\alpha\beta\bar{r} \leqslant 1$, entonces

$$\|y_{k+1} - u\| \leqslant \alpha\beta\bar{r} \|y_k - u\| \leqslant \|y_k - u\| \leqslant \bar{r}.$$

Por lo tanto, de la condición $y_k \in \Omega(\bar{r})$ se sigue que $y_{k+1} \in \Omega(\bar{r})$. Ya que $y_0 \in \Omega(\bar{r})$, entonces por inducción encontramos que $y_k \in \Omega(\bar{r})$ para cualquier $k \geqslant 0$. Por eso la estimación (24) es válida para cualquier $k \geqslant 0$.

Resolvamos la desigualdad (24). Multipliquémosla por $\alpha\beta$ y designemos $q_k = \alpha\beta \|y_k - u\|$. Para q_k obtendremos la desigualdad $q_{k+1} \leqslant q_k^2$, $k = 0, 1, \ldots$ Por inducción

es fácil demostrar que su solución tiene la forma $q_n \leqslant q_0^{2^n}$, $n \geqslant 0$. Por consiguiente, tenemos la estimación

$$\alpha\beta \, \| y_n - u \| \leqslant (\alpha\beta \, \| y_0 - u \|)^{2^n}.$$

De aquí se deduce la afirmación del teorema.

OBSERVACION 1. *Si la aproximación inicial y_0 es elegida de manera tal, que $\bar{r} \leqslant \rho/(\alpha\beta)$, $\rho < 1$, entonces de* (23) *se deduce la estimación*

$$\|y_n - u\| \leqslant \rho^{2^n - 1} \|y_0 - u\|$$

y la estimación

$$n \geqslant n_0 (\varepsilon) = \log_2 (\ln \varepsilon / \ln \rho + 1)$$

para el número de iteraciones.

OBSERVACION 2. *Si en lugar de la condición* (21) *se cumple la desigualdad*

$$\| A' (v) - A' (w) \| \leqslant \alpha \| v - w \|^p, \quad p \in (0, 1),$$

entonces para el error es cierta la estimación

$$\|y_n - u\| \leqslant \frac{1}{\sqrt[p]{\alpha\beta}} (\sqrt[p]{\alpha\beta} \, \|y_0 - u\|)^{(p+1)^n},$$

$$\left(\tilde{r} = \min \left(r, \frac{1}{\sqrt[p]{\alpha\beta}} \right)\right).$$

Durante la demostración del teorema 4 nosotros obtuvimos la estimación para el error (23). Esta estimación es inútil desde el punto de vista de su aplicación práctica, pero es importante para la teoría del método, por cuanto muestra cómo se realiza la convergencia cerca de la solución u.

El teorema 4 permite encontrar el dominio de no existencia de solución. En efecto, el teorema afirma que y_k converge a u si $\| y_0 - u \| \leqslant \bar{r}$. Por lo tanto, si las iteraciones no convergen, entonces en la esfera $\| y_0 - u \| \leqslant \bar{r}$, con centro en el punto y_0 no hay soluciones de la ecuación (1).

Señalemos que si el operador A tiene segunda derivada de Gato en la esfera $\Omega (r)$, entonces en la desigualdad (21)

$$\alpha = \sup_{0 \leqslant t \leqslant 1} \|A'' [(v + t (w - v))]\|.$$

Al realizar el esquema iterativo (19) para cada k es necesario resolver la ecuación operacional lineal

$$A' (y_k) v = F (y_k), \tag{25}$$

donde

$$F(y_k) = A'(y_k)\, y_k - \tau_{k+1}(Ay_k - f). \qquad (26)$$

Si v es la solución exacta de la ecuación (25), entonces en (19) $y_{k+1} = v$.

Es necesario calcular el operador $A'(y_k)$ en cada iteración y esto puede exigir grandes gastos computacionales. Examinemos un ejemplo. Supongamos que el operador A corresponde al sistema de ecuaciones no lineales

$$\varphi_i(u) = 0, \quad i = 1, 2, \ldots, m, \quad u = (u_1, u_2, \ldots, u_m).$$

La derivada de Gato $A'(y)$ en el punto $y = (y_1, y_2, \ldots, y_m)$ es una matriz cuadrada con elementos $a_{ij}(y)$, donde

$$a_{ij}(y) = \frac{\partial \varphi_i(u)}{\partial u_j}\Big|_{u=y}, \quad i, j = 1, 2, \ldots, m.$$

Por lo tanto, en cada iteración es necesario calcular m^2 elementos de la matriz $A'(y)$, al mismo tiempo que el número de incógnitas en el problema es igual a m.

Para evitar el cálculo de la derivada $A'(y_k)$ en cada iteración, se utiliza la siguiente modificación del esquema (19):

$$A'(y_{km})\frac{x_{km+i+1} - y_{km+i}}{\tau_{km+i+1}} + Ay_{km+i} = f,$$

$$i = 0, 1, \ldots, m-1, \quad k = 0, 1, \ldots$$

Aquí la derivada A' se calcula después de cada m iteraciones y se utiliza para encontrar las aproximaciones intermedias $y_{km+1}, y_{km+2}, \ldots, y_{(k+1)m}$. Para $m = 1$ obtendremos el esquema iterativo (19).

4. Métodos iterativos bietápicos. Es racional utilizar el esquema iterativo (19) en el caso cuando el operador $A'(y_k)$ es fácilmente inversible. Con esto la solución exacta v de la ecuación (25) se toma por una nueva aproximación iterativa y_{k+1}, la cual satisface el esquema (19). De este modo, nosotros obtenemos un esquema iterativo, en el cual el operador B_{k+1} está prefijado en forma explícita: $B_{k+1} = A'(y_k)$.

Si la ecuación (25) se resuelve aproximadamente, por ejemplo, con ayuda de un método iterativo (interno) auxiliar y en calidad de y_{k+1} se toma la m-ésima aproximación iterativa v_m, entonces y_{k+1} satisface el esquema general (18)

con algún $B_{k+1} \neq A'(y_k)$. En este caso la forma explícita del operador B_{k+1} no se utiliza, y el conocimiento de su estructura es necesario únicamente para investigar la convergencia del esquema iterativo (18). Los métodos iterativos construidos por este procedimiento en algunas ocasiones son llamados bietápicos, sobreentendiendo por esto el algoritmo especial de inversión del operador B_{k+1}.

Describamos más detalladamente el esquema general de construcción de los métodos bietápicos. Supongamos que para resolver la ecuación lineal (25) se utiliza algún método iterativo de dos capas implícito

$$\overline{B}_{n+1} \frac{v_{n+1} - v_n}{\omega_{n+1}} + A'(y_k) v_n = F(y_k), \quad n = 0, 1, \ldots, m-1, \tag{27}$$

donde $F(y_k)$ está definido en (26), $\{\omega_n\}$ es el conjunto de los parámetros iterativos, $\overline{B}_{n+1}$ son los operadores en H que pueden depender de y_k, y $v_0 = y_k$.

Expresemos v_m mediante y_k. Primeramente hallaremos la ecuación para el error $z_n = v_n - v$, donde v es la solución de la ecuación (25). De (25) y (27) hallaremos

$$z_{n+1} = S_{n+1} z_n, \quad n = 0, 1, \ldots, \quad S_n = E - \omega_n \overline{B}_n^{-1} A'(y_k)$$

y, por consiguiente,

$$z_m = v_m - v = T_m z_0 = T_m (v_0 - v), \quad T_m = S_m S_{m-1} \ldots, S_1, \tag{28}$$

$$v_m = (E - T_m) v + T_m y_k.$$

De (25) y (26) obtendremos

$$v = [A'(y_k)]^{-1} F(y_k) = y_k - \tau_{k+1} [A'(y_k)]^{-1} (A y_k - f).$$

Sustituyendo el v hallado en (28), tendremos

$$y_{k+1} = v_m = y_k - \tau_{k+1} (E - T_m) [A'(y_k)]^{-1} (A y_k - f).$$

De aquí se deduce que y_{k+1} satisface el esquema iterativo (18), si designamos

$$B_{k+1} = A'(y_k) (E - T_m)^{-1}. \tag{29}$$

Así pues, la realización de un paso del método bietápico consiste en el cálculo de $F(y_k)$ por la fórmula (26) y el cumplimiento de m iteraciones según el esquema (27) con la

aproximación inicial $v_0 = y_k$. La aproximación v_m obtenida se toma en calidad de y_{k+1}.

Examinemos el esquema iterativo (18), (19). Para estimar la velocidad de convergencia se pueden utilizar los teoremas 2 y 3, en los cuales B se cambia por B_{k+1}, y τ por τ_{k+1}. La insuficiencia de esta elección del parámetro τ es que se necesita estimar γ_1, γ_2 y γ_3 exactamente.

Señalemos que durante la construcción de un método bietápico se podría partir no de la ecuación (25), sino de la ecuación «cercana» a ella

$$Rv = F(y_k),$$

donde el operador lineal R en cierto sentido es equivalente al operador $A'(y_k)$. En este caso en el esquema iterativo (18) tenemos

$$B_{k+1} \equiv B = R(E - T_m)^{-1}.$$

Investiguemos este caso más detalladamente. Supongamos cumplidas las condiciones

$$R = R^* > 0, \quad T_m^* R = RT_m, \tag{30}$$

$$\| T_m \|_R \leqslant q < 1. \tag{31}$$

LEMA 2. *Sean cumplidas las condiciones* (30) *y* (31). *Entonces el operador* $B = R(E - T_m)^{-1}$ *es autoconjugado y definido positivo en* H, *y son válidas las desigualdades*

$$(1 - q) B \leqslant R \leqslant (1 + q) B. \tag{32}$$

Examinemos el operador $B^{-1} = (E - T_m) R^{-1}$. De (30) hallaremos $(E - T_m^*) R = R(E - T_m)$ o $R^{-1}(E - T_m^*) = (E - T_m) R^{-1}$. Por consiguiente, el operador B^{-1} es autoconjugado en H.

Puesto que en virtud de (30) el operador T_m es autoconjugado en H_R, entonces

$$\|T_m\|_R = \sup_{x \neq 0} \frac{|(T_m x, x)_R|}{(x, x)_R} = \sup_{x \neq 0} \frac{|(RT_m x, x)|}{(Rx, x)} \leqslant q < 1.$$

Por lo tanto, para cualquier $x \in H$ tenemos la desigualdad

$$|(RT_m x, x)| \leqslant q (Rx, x).$$

Poniendo aquí $x = R^{-1}y$, obtendremos

$$|(T_m R^{-1} y, y)| \leqslant q (R^{-1} y, y),$$

por lo tanto para $y \in H$ hallaremos

$$(1 - q)\,(R^{-1}y,\ y) \leqslant ((E - T_m)\,R^{-1}y,\ y) \leqslant (1 + q) \times (R^{-1}y,\ y).$$

Así pues, hemos obtenido la estimación

$$(1 - q)\,R^{-1} \leqslant B^{-1} \leqslant (1 + q)\,R^{-1}. \tag{33}$$

Ya que R^{-1} y B^{-1} son los operadores autoconjugados en H y $q < 1$, entonces del lema 9 del § 1 del cap. V se deduce que las desigualdades (33) y (32) son equivalentes. El lema está demostrado.

LEMA 3. *Supongamos que el operador A tiene en la esfera $\Omega(r)$ derivada de Gato $A'(v)$, la cual para cualquier $v \in \in \Omega(r)$ satisface las desigualdades*

$$c_1\,(Ry,\ y) \leqslant (A'(v)\,y,\ y) \leqslant c_2\,(Ry,\ y), \quad c_1 > 0, \tag{34}$$

$$\|0{,}5\,[A'(v) - (A'(v))^*]\,y\|^2_{R^{-1}} \leqslant c_3^2\,(Ry,\ y), \quad c_3 \geqslant 0. \tag{35}$$

y sean cumplidas las condiciones (30) *y* (31). *Entonces se cumplen las desigualdades* (17) *del teorema* 3, *donde*

$$\bar{\gamma}_1 = c_1\,(1 - q),\ \bar{\gamma}_2 = c_2\,(1 + q),\ \bar{\gamma}_3 = c_3\,(1 + q)^2,$$
$$D = B = R\,(E = T_m).$$

En efecto, en virtud del lema 2 el operador D es autoconjugado y definido positivo en H. Además, las desigualdades (17) para $D = B$ tienen la forma

$$\bar{\gamma}_1\,(By,\ y) \leqslant (A'(v)\,y,\ y) \leqslant \bar{\gamma}_2\,(By,\ y), \tag{36}$$

$$\|0{,}5\,[A'(v) - A'(v))^*]\,y\|^2_{B^{-1}} \leqslant \bar{\gamma}_3^2 (By, y). \tag{37}$$

Las desigualdades (36) con los $\bar{\gamma}_1$ y $\bar{\gamma}_2$ indicados en el lema 3 se deducen de (32) y (34), mientras (37) se desprende de (32), (33) y (35), ya que

$$\|z\|^2_{B^{-1}} = (B^{-1}z,\ z) \leqslant (1 + q)\,(R^{-1}z,\ z) = (1 + q)\|z\|^2_{R^{-1}},$$
$$(Rz,\ z) \leqslant (1 + q)\,(Bz,\ z).$$

Utilizando el lema 3 se puede demostrar el análogo del teorema 3 para el método bietápico.

TEOREMA 5. *Sean cumplidas las condiciones del lema* 3 *y sea el método bietápico construido a base de la ecuación $Rv = \\ = F(y_k)$ empleando el operador de resolución T_m. Si en el esquema iterativo* (18) *con $B_{k+1} \equiv B - R\,(E - T_m)^{-1}$, que pescribe este método bietápico, elegimos $\tau_k \equiv \tau_0\,(1 - \varkappa\bar{\rho})$*

e $y_0 \in \Omega(r)$, *entonces para el error será cierta la estimación*

$$\|y_n - u\|_B \leqslant \bar{\rho}^n \|y_0 - u\|_B,$$

donde u *es la solución de la ecuación* (1) *y* $\bar{\rho}$, $\varkappa$ *y* τ_0 *son definidos en* (16) *con los* $\bar{\gamma}_1$, $\bar{\gamma}_2$ *y* $\bar{\gamma}_3$ *indicados en el lema* 3.

5. Otros métodos iterativos. En este punto nosotros daremos una breve descripción de algunos métodos iterativos también utilizados para resolver la ecuación del operador A no lineal.

Sea $\Phi(u)$ un funcional diferenciable según Gato, definido en H. El operador A, que actúa en H, se llama *potencial*, si existe el funcional diferenciable $\Phi(u)$ tal, que $Au = \operatorname{grad} \Phi(u)$ para todos los u. Aquí el gradiente del funcional $\Phi(u)$ se define por la igualdad $\frac{d}{dt}\Phi(u + tv)|_{t=0} = (\operatorname{grad} \Phi(u), v)$. Como ejemplo del operador potencial puede servir un operador lineal A acotado y autoconjugado en el espacio de Hilbert H. Él está generado por el funcional $\Phi(u) = 0{,}5\,(Au, u)$.

Sea el operador A continuamente diferenciable en H. El operador A es potencial entonces, y solamente entonces, cuando la derivada de Gato $A'(v)$ es un operador autoconjugado en H.

Si el operador A es potencial, entonces la fórmula

$$\Phi(u) = \int_0^1 (A(u_0 + t(u - u_0)),\ u - u_0)\, dt,$$

donde u_0 es arbitrario, pero un elemento fijo de H, da un procedimiento para construir el funcional $\Phi(u)$ según el operador A.

Si el operador A está generado por el gradiente de un funcional estrictamente convexo, entonces la derivada $A'(v)$ es un operador definido positivo en H para cualquier $v \in H$. En este caso para la resolución aproximada de la ecuación (25) se pueden utilizar los métodos iterativos de tipo variacional, por ejemplo, en (27) se pueden elegir los parámetros ω_{n+1} por las fórmulas de los métodos del descenso más rápido, de los defectos mínimos, etc.

Examinemos como ejemplo el método bietápico (18), (29), para el cual $\tau_{k+1} \equiv 1$, y en el esquema (27) $m = 1$ y $B_1 = E$. Entonces $B_{k+1} = E/\omega_1$. Si para el proceso iterativo auxiliar (27) se elige el parámetro ω_1 por las fórmulas

del método de los defectos mínimos (o de las correcciones mínimas), entonces obtendremos (véanse los puntos 2, 3 del § 2 del cap. VIII)

$$\omega_1 = \frac{(A'(y_k) r_k, r_k)}{\|A'(y_k) r_k\|^2}, \quad r_k = Ay_k - f. \tag{38}$$

En este caso el método iterativo bietápico se describe por la fórmula

$$\frac{y_{k+1} - y_k}{\omega_1} + Ay_k = f, \quad k = 0, 1, \ldots, \tag{39}$$

donde ω_1 está definido en (38).

En la situación, cuando el operador A no es potencial, el parámetro ω_1 se puede elegir mediante las fórmulas del método de los errores mínimos, poniendo en (27) $B_1 = [(A'(y_k))^*]^{-1}$ y

$$\omega_1 = \frac{(r_k, r_k)}{\|A'(y_k))^* r_k\|^2}, \quad r_k = Ay_k - f. \tag{40}$$

En este caso el método bietápico tiene la forma

$$\frac{y_{k+1} - y_k}{\omega_1} + (A'_*(y_k))^* Ay_k = (A'(y_k))^* f, \quad k = 0, 1, \ldots, \tag{41}$$

donde ω_1 está definido en (40).

Es fácil ver, que en el método (38), (39) el parámetro ω_1 se elige de la condición de mínimo de $\| A'(y_k)(y_{k+1} - y_k) + Ay_k - f \|$, y en el método (40), (41), de la condición de mínimo de la norma $\| y_{k+1} - y_k + [A'(y_k)]^{-1} \times (Ay_k - f) \|$.

El problema de resolver la ecuación $Au = f$ en el caso de un operador potencial algunas veces puede ser cambiado por el problema de minimizar el funcional que genera este operador. Notemos que siempre se tiene un procedimiento simple de transformar el problema de resolver la ecuación (1) en un problema de minimización aún si el operador A no es potencial.

En efecto, sea $\Phi(u)$ un funcional definido en H y que tiene el único punto de mínimo $u = 0$. En calidad de ejemplo de tal funcional se puede citar $\Phi(u) = (Du, u)$, donde D es un operador autoconjugado y definido positivo en H. A continuación, para la ecuación (1) dada examinemos el funcional

$$F(u) = \Phi(Au - f), \quad u \in H.$$

Si la ecuación (1) tiene solución u, entonces, evidentemente ella le concede el mínimo al funcional $F(u)$.

Describamos un método de minimización de un funcional (método del descenso). Sea la ecuación (1) generada por el gradiente de un funcional $\Phi(u)$ estrictamente convexo. Sea la sucesión minimizante se construye de acuerdo con el esquema iterativo (19), es decir, mediante la fórmula

$$y_{k+1} = y_k - \tau_{k+1} [A'(y_k)]^{-1} (Ay_k - f), \quad k = 0, 1, \ldots \quad (42)$$

Designemos

$$w_k = [A'(y_k)]^{-1} \operatorname{grad} \Phi(y_k),$$

donde en virtud de las suposiciones hechas $\operatorname{grad} \Phi(y_k) = Ay_k - f$. Escribamos (42) en la forma

$$y_{k+1} = y_k - \tau_{k+1} w_k.$$

Señalemos, que el operador $A'(y_k)$ es definido positivo y autoconjugado en H. A continuación, de la definición de la derivada de Gato de un funcional tenemos

$$\lim_{\tau_{k+1} \to 0} \left[\frac{\Phi(y_k - \tau_{k+1} w_k) - \Phi(y_k)}{\tau_{k+1}} + (\operatorname{grad} \Phi(y_k), w_k) = 0. \right.$$

Ya que $A'(y_k) w_k = \operatorname{grad} \Phi(y_k)$, entonces

$$(\operatorname{grad} \Phi(y_k), w_k) = (A'(y_k) w_k, w_k) > 0.$$

Por consiguiente, existe un $\tau_{k+1} > 0$, tal, que $\Phi(y_{k+1})$ será estrictamente menor que $\Phi(y_k)$.

Si la sucesión minimizante $\{y_k\}$ se construye según el esquema explícito (18) ($B_k \equiv E$), es decir, mediante las fórmulas

$$y_{k+1} = y_k - \tau_{k+1} (Ay_k - f),$$

entonces el paso de y_k a y_{k+1} se realiza por la dirección del gradiente del funcional $\Phi(u)$ en el punto y_k. Tales métodos se les acepta llamar *métodos de descenso gradiental*. Existen algunos algoritmos de elección de los parámetros iterativos τ_k, pero en estos problemas nosotros no nos detendremos aquí.

Citemos como conclusión una generalización del método explícito de los gradientes conjugados, el cual se utiliza para minimizar un funcional bajo las suposiciones indicadas más arriba. Las fórmulas del algoritmo d eFletcher—

Ribs tienen la forma:

$$y_{k+1} = y_k - a_{k+1} w_k, \quad k = 0, 1, \ldots,$$
$$w_k = \operatorname{grad} \Phi(y_k) + b_k w_{k-1}, \quad k = 1, 2, \ldots,$$
$$w_0 = \operatorname{grad} \Phi(y_0),$$

donde

$$b_k = \frac{\| \operatorname{grad} \Phi(y_k) \|^2}{\| \operatorname{grad} \Phi(y_{k-1}) \|^2}, \quad k = 1, 2, \ldots,$$

y el parámetro a_{k+1} se elige de la condición de mínimo de $\Phi(y_k - a_{k+1} w_k)$. Este problema sobre la búsqueda del mínimo de función de una variable se resuelve por uno de los métodos del análisis numérico.

§ 2. Métodos de resolución de esquemas de diferencias no lineales

1. Esquema de diferencias para la ecuación cuasilineal elíptica unidimensional. Aplicaremos la teoría general de los esquemas iterativos expuesta en el § 1 para encontrar una solución aproximada de los esquemas de diferencias elípticos no lineales. Comencemos por los ejemplos más simples.

Examinemos el tercer problema de contorno para una ecuación cuasilineal unidimensional en la forma divergente

$$Lu = \frac{d}{dx} k_1\left(x, u, \frac{du}{dx}\right) - k_0\left(x, u, \frac{du}{dx}\right) = -\varphi(x), \quad 0 \leqslant x \leqslant l,$$

$$k_1\left(x, u, \frac{du}{dx}\right) = \varkappa_0(u) - \mu_0, \quad x = 0, \tag{1}$$

$$-k_1\left(x, u, \frac{du}{dx}\right) = \varkappa_1(u) - \mu_1, \quad x = l.$$

Supondremos que las funciones $k_1(x, p_0, p_1)$, $k_0(x, p_0, p_1)$, $\varkappa_0(p_0)$ y $\varkappa_1(p_0)$ sean continuas respecto de p_0 y p_1 y estén cumplidas las condiciones de elipticidad fuerte

$$\sum_{\alpha=0}^{1} [k_\alpha(x, p_0, p_1) - k_\alpha(x, q_0, q_1)](p_\alpha - q_\alpha) \geqslant$$
$$\geqslant c_1 \sum_{\alpha=0}^{1} (p_\alpha - q_\alpha)^2, \tag{2}$$

$$[\varkappa_\alpha(p_0) - \varkappa_\alpha(q_0)](p_0 - q_0) \geqslant 0, \quad \alpha = 0, 1, \tag{3}$$

donde $c_1 > 0$ es la constante positiva, $0 \leqslant x \leqslant l$, $|p_0|$, $|q_0|$, $|p_1|$, $|q_1| < \infty$.

Sobre la red uniforme $\bar{\omega} = \{x_i = ih,\ i = 0, 1, \ldots, N,\ hN = l\}$ al problema (1) le pondremos en correspondencia el esquema de diferencias

$$\Lambda y_i = -f_i, \quad 0 \leqslant i \leqslant N, \tag{4}$$

donde

$$f_i = \begin{cases} \varphi(0) + \frac{2}{h}\mu_0, & i = 0, \\ \varphi(x_i), & 1 \leqslant i \leqslant N-1, \\ \varphi(l) + \frac{2}{h}\mu_1, & i = N. \end{cases}$$

El operador de diferencias Λ se define por las fórmulas:

$$\Lambda y_i = 1/2\{[k_1(x, y, y_x)]_{\bar{x}} + [k_1(x, y, y_{\bar{x}})]_x - k_0(x, y, y_x) - k_0(x, y, y_{\bar{x}})\}_i, \quad 1 \leqslant i \leqslant N-1,$$

$$\Lambda y_0 = \frac{1}{h}[k_1(0, y_0, y_{x,0}) + k_1(h, y_1, y_{\bar{x},1})] - k_0(0, y_0, y_{x,0}) - \frac{2}{h}\varkappa_0(y_0), \quad i = 0,$$

$$\Lambda y_N = -\frac{1}{h}[k_1(l-h, y_{N-1}, y_{x,N-1}) + k_1(l, y_N, y_{\bar{x},N})] - k_0(l, y_N, y_{\bar{x},N}) - \frac{2}{h}\varkappa_1(y_N), \quad i = N.$$

Si en el espacio $H = H(\bar{\omega})$ se define el operador no lineal A por la relación $A = -\Lambda$, entonces el esquema de diferencias (4) se escribe en la forma de la ecuación operacional $Au = f$.

Investiguemos las propiedades del operador no lineal A, el cual actúa de H en H. Recordemos que el producto escalar en $H(\bar{\omega})$ se define por la fórmula

$$(u, v) = \sum_{i=1}^{N-1} u_i v_i h + 0{,}5h(u_0 v_0 + u_N v_N),$$

y mediante $(u, v)_{\omega+}$ y $(u, v)_{\omega-}$ se designan las sumas

$$(u, v)_{\omega+} = \sum_{i=1}^{N} u_i v_i h, \quad (u, v)_{\omega-} = \sum_{i=0}^{N-1} u_i v_i h,$$

así que

$$(u,\ v) = {}^1/_2\,[(u,\ v)_{\omega+} + (u,\ v)_{\omega-}].$$

Mostremos que al cumplirse las condiciones (2) y (3) el operador A es fuertemente monótono en $H(\overline{\omega})$, es decir, se cumple la desigualdad

$$(Au - Av,\ u - v) \geqslant c_1 \| u - v \|^2, \quad c_1 > 0, \qquad (5)$$

donde c_1 está definido en (2).

Designemos $p_0 = \overline{p}_0 = u_t$, $q_0 = \overline{q}_0 = v_t$, $p_1 = u_{x,\,t}$, $\overline{p}_1 = u_{\overline{x},\,t}$, $q_1 = v_{x,\,t}$, $q_1 = v_{\overline{x},\,t}$.

Utilizando la definición del operador A, las fórmulas de sumación por partes (véase (7), (9) del § 2 del cap. V) y las condiciones (2), (3), obtendremos

$$(Au - Av,\ u - v) = (\Lambda v - \Lambda u,\ u - v) =$$

$$= {}^1/_2 \sum_{i=1}^{N} h \left\{ \sum_{\alpha=0}^{1} [k_\alpha (x,\ \overline{p}_0,\ \overline{p}_1) - \right.$$

$$\left. - k_\alpha (x,\ \overline{q}_0,\ \overline{q}_1)]\,(\overline{p}_\alpha - \overline{q}_\alpha) \right\}_i +$$

$$+ {}^1/_2 \sum_{i=0}^{N-1} h \left\{ \sum_{\alpha=0}^{1} [k_\alpha (x,\ p_0,\ p_1) - k_\alpha (x,\ q_0,\ q_1)] \times \right.$$

$$\left. \times (p_\alpha - q_\alpha) \right\}_i + [\varkappa_1 (\overline{p}_0) - \varkappa_1 (\overline{q}_0)]\,(\overline{p}_0 - \overline{q}_0) \Big|_{i=N} +$$

$$+ [\varkappa_0 (p_0) - \varkappa_0 (q_0)]\,(p_0 - q_0)\,|_{i=0} \geqslant$$

$$\geqslant \frac{c_1}{2} \sum_{i=1}^{N} h \sum_{\alpha=0}^{1} (\overline{p}_\alpha - \overline{q}_\alpha)_i^2 + \frac{c_1}{2} \sum_{i=0}^{N-1} h \sum_{\alpha=0}^{1} (p_\alpha - q_\alpha)_i^2.$$

Teniendo en cuenta la igualdad $u_{x,\,t} = u_{\overline{x},\,t+1}$, escribamos la estimación obtenida en la forma

$$(Au - Av,\ u - v) \geqslant \frac{c_1}{2} [u - v,\ u - v)_{\omega+} + (u - v,\ u - v)_{\omega-} +$$

$$+ \sum_{i=1}^{N} h\,(u - v)_{x,\,i}^2 \sum_{i=0}^{N-1} h\,(u - v)_{\overline{x},\,i}^2] = c_1 [\|u - v\|^2 +$$

$$+ ((u - v)_{\overline{x}}^2,\ 1)_{\omega+}] \geqslant c_1 \|u - v\|^2.$$

De la observación 2 al lema 12 del capítulo V se deduce que esta estimación no puede ser mejorada.

Así, pues, está establecida la monotonía fuerte del operador A. En virtud de la continuidad de las funciones $k_\alpha(x, p_0, p_1)$ y $\varkappa_\alpha(p_0)$, el operador A es continuo en H. Por eso del teorema 11 del capítulo V se deduce la existencia y la unicidad de solución de la ecuación $Au = f$ en la esfera $\| u \| \leqslant \frac{1}{c_1} \| A0 - f \|$ y, por consiguiente, del problema de diferencias (4).

Si $k_\alpha(x, p_0, p_1)$ y $\varkappa_\alpha(p_0)$, $\alpha = 0, 1$ son las funciones continuamente diferenciables de sus argumentos, entonces en lugar de (2) y (3) se pueden utilizar otras condiciones suficientes que aseguran la monotonía fuerte del operador A.

Supondremos que se cumplan las condiciones

$$c_1 \sum_{\alpha=1}^{1} \xi_\alpha^2 \leqslant \sum_{\alpha,\beta=0}^{1} a_{\alpha\beta}(x, p_0, p_1)\,\xi_\alpha\xi_\beta \leqslant c_2 \sum_{\alpha=0}^{1} \xi_\alpha^2, \quad c_1 > 0, \tag{6}$$

$$0 \leqslant \sigma_\alpha(p_0) \leqslant c_3, \quad \alpha = 1, 2, \tag{7}$$

donde $\xi = (\xi_0, \xi_1)$ es un vector arbitrario y

$$a_{\alpha\beta}(x, p_0, p_1) = \frac{\partial k_\alpha(x, p_0, p_1)}{\partial p\beta},$$

$$\sigma_\alpha(p_0) = \frac{\partial \varkappa_\alpha(p_0)}{\partial p_0}, \quad \alpha, \beta = 0, 1.$$

Mostremos, que de las condiciones (6) y (7) se deducen (2) y (3). En realidad, tienen lugar las igualdades

$$k_\alpha(x, p_0, p_1) - k_\alpha(x, q_0, q_1) =$$

$$= \int_0^1 \frac{d}{dt} k_\alpha(x, tp_0 + (1-t)q_0, tp_1 + (1-t)q_1)\,dt =$$

$$= (p_0 - q_0) \int_0^1 \frac{\partial k_\alpha(x, s_0, s_1)}{\partial s_0}\,dt + (p_1 - q_1) \times$$

$$\times \int_0^1 \frac{\partial k_\alpha(x, s_0, s_1)}{\partial s_1}\,dt = \sum_{\beta=0}^{1} (p_\beta - q_\beta) \times$$

$$\times \int_0^1 a_{\alpha\beta}(x, s_0, s_1)\,dt, \quad \alpha = 0, 1,$$

donde $s_0 = tp_0 + (1-t)q_0$, $s_1 = tp_1 + (1-t)q_1$. Multiplicando esta igualdad por $p_\alpha - q_\alpha$ y sumándola según

α desde 0 hasta 1, obtendremos teniendo en cuenta (6)

$$\sum_{\alpha=0}^{1} [k_\alpha(x,\ p_0,\ p_1) - k_\alpha(x,\ q_0,\ q_1)]\,(p_\alpha - q_\alpha) =$$

$$= \int_0^1 \sum_{\alpha,\,\beta=0}^{1} a_{\alpha\beta}(x,\ s_0,\ s_1)\,(p_\alpha - q_\alpha)\,(p_\beta - q_\beta)\,dt \geqslant$$

$$\geqslant c_1 \int_0^1 \sum_{\alpha=0}^{1} (p_\alpha - q_\alpha)^2\,dt = c_1 \sum_{\alpha=0}^{1} (p_\alpha - q_\alpha)^2.$$

Así, hemos obtenido la desigualdad (2). Del modo análogo de (7) obtenemos la desigualdad (3)

$$[\varkappa_\alpha(p_0) - \varkappa_\alpha(q_0)]\,(p_0 - q_0) = \int_0^1 \frac{\partial \varkappa_\alpha(s_0)}{\partial s_0}\,dt\,(p_0 - q_0)^2 \geqslant 0.$$

De esta forma, las condiciones (6) y (7) garantizan la existencia y la unicidad de solución del problema de diferencias (4).

Hallemos ahora la derivada de Gato del operador A, suponiendo que las funciones $k_\alpha(x,\ p_0,\ p_1)$ y $\varkappa_\alpha(p_0)$, $\alpha = 0,\ 1$, tienen derivadas acotadas según p_0 y p_1 del orden necesitado.

De la definición de la derivada de Gato de un operador no lineal tendremos

$$A'(u)\,y_i = -{}^1/_2 \{[a_{11}(x,\ u,\ u_x)\,y_x]^-_{x,\,i} +$$

$$+ [a_{11}(x,\ u,\ u^-_{\bar{x}})\,y^-_{\bar{x}}]_{x,\,i} + [a_{10}(x,\ u,\ u_x)\,y]^-_{x,\,i} +$$

$$+ [a_{10}(x,\ u,\ u^-_{\bar{x}})\,y]_{x,\,i}\} + {}^1/_2 \{a_{01}(x,\ u,\ u_x)\,y_{x,\,i} +$$

$$+ a_{01}(x,\ u,\ u^-_{\bar{x}})\,y^-_{\bar{x},\,i} + [a_{00}(x,\ u,\ u_x) +$$

$$+ a_{00}(x,\ u,\ u^-_{\bar{x}})]\,y_i\},\quad 1 \leqslant i \leqslant N-1.$$

Para $i = 0$ obtendremos

$$A'(u)\,y_0 = -\frac{1}{h}[a_{11}(0,\ u_0,\ u_{x,\,0}) + a_{11}(h,\ u_1,\ u^-_{\bar{x},\,1}) -$$

$$- h a_{01}(0,\ u_0,\ u_{x,\,0}) + h a_{10}(h,\ u_1,\ u^-_{\bar{x},\,1})]\,y_{x,\,0} +$$

$$+ \frac{2}{h}[\sigma_0(u_0) - {}^1/_2\,a_{10}(0,\ u_0,\ u_{x,\,0}) -$$

$$- {}^1/_2\,a_{10}(h,\ u_1,\ u^-_{\bar{x},\,1}) + \frac{h}{2}\,a_{00}(0,\ u_0,\ u_{x,\,0})]\,y_0,$$

y para $i = N$ tendremos

$$A'(u)\,y_N = \frac{1}{h}\,[a_{11}(l-h,\ u_{N-1},\ u_{x,\,N-1}) +$$
$$+ a_{11}(l,\ u_N,\ u_{\bar{x},\,N}) + h a_{01}(l,\ u_N,\ u_{\bar{x},\,N}) -$$
$$- h a_{10}(l-h,\ u_{N-1},\ u_{x,\,N-1})]\,y_{\bar{x},\,N} +$$
$$+ \frac{2}{h}\,[\sigma_1(u_N) + {}^1/_2\,a_{10}(l,\ u_N,\ u_{\bar{x},\,N}) +$$
$$+ {}^1/_2\,a_{10}(l-h,\ u_{N-1},\ u_{x,\,N-1}) + \frac{h}{2}\,a_{00}(l,\ u_N,\ u_{\bar{x},\,N})]\,y_N.$$

Señalemos que al calcular $A'(u)\,y_0$ y $A'(u)\,y_N$ fueron utilizadas las relaciones

$$y_1 = y_0 + h y_{x,\,0}, \quad y_{N-1} = y_N - h y_{\bar{x},N}. \tag{8}$$

Investiguemos las propiedades de la derivada de Gato $A'(u)$ del operador A.

Lema 4. *Si son cumplidas las condiciones*

$$\frac{\partial k_1(x,\ p_0,\ p_1)}{\partial p_0} = \frac{\partial k_0(x,\ p_0,\ p_1)}{\partial p_1}, \tag{9}$$

entonces $A'(u)$ es un operador autoconjugado en H. Al satisfacer las condiciones (6) y (7) él es definido positivo en H.

En efecto, utilizando las fórmulas de sumación por partes, y también las relaciones (8), obtenemos

$$(A'_1(u)\,y,\ z) = {}^1/_2 \sum_{i=0}^{N-1} h\,[a_{11}(x,\ u,\ u_x)\,y_x z_x +$$
$$+ a_{10}(x,\ u,\ u_x)\,y z_x + a_{01}(x,\ u,\ u_x)\,y_x z +$$
$$+ a_{00}(x,\ u,\ u_x)\,yz]_i + {}^1/_2 \sum_{i=1}^{N} h\,[a_{11}(x,\ u,\ u_{\bar{x}})\,y_{\bar{x}} z_{\bar{x}} +$$
$$+ a_{10}(x,\ u,\ u_{\bar{x}})\,y z_{\bar{x}} + a_{01}(x,\ u,\ u_{\bar{x}})\,y_{\bar{x}} z +$$
$$+ a_{00}(x,\ u,\ u_x)\,yz]_i + \sigma_0(u_0)\,y_0 z_0 + \sigma_1(u_N)\,y_N z_N. \tag{10}$$

Comparando esta expresión con la expresión para $(y,\ A'(u)\,z)$ obtendremos que bajo la condición $a_{10}(x,\ p_0,\ p_1) = a_{01}(x,\ p_0,\ p_1)$, la cual es otra forma de escritura para (9), el operador $A'(u)$ es autoconjugado en H para cualquier $u \in H$.

Sean ahora cumplidas las condiciones (6) y (7). Poniendo en (10) $z_i \equiv y_i$, obtendremos

$$(A'(u)\,y,\ y) \geqslant \frac{c_1}{2}\Big[\sum_{i=0}^{N-1} h\,(y_i^2 + y_{\overline{x},\,i}^2) + \\ + \sum_{i=1}^{N} h\,(y_i^2 + y_{\overline{x},\,i}^2)\Big] = c_1\,[(y,\ y) + (y_{\overline{x}}^2,\ 1)_{\omega+}] \geqslant c_1\,(y,\ y), \quad (11)$$

es decir, el operador $A'(u)$ es definido positivo en H. El lema está demostrado.

Notemos que en virtud del teorema 2 del capítulo V de la definición positiva de la derivada de Gato del operador continuo A se deduce que él es fuertemente monótono. De esta forma, al cumplirse las condiciones (6), (7) el operador A es fuertemente monótono.

Poniendo $z_i \equiv y_i$ en (10), obtenemos en virtud de las condiciones (6), (7) la estimación por encima

$$(A'(u)\,y,\ y) \leqslant \frac{c_2}{2}\Big[\sum_{i=0}^{N-1} h\,(y_i^2 + y_{\overline{x},\,i}^2) + \sum_{i=1}^{N} h\,(y_i^2 + y_{\overline{x},\,i}^2)\Big] + \\ + c_3\,(y_0^2 + y_N^2) = c_2\,[(y,\ y) + (y_{\overline{x}}^2,\ 1)_{\omega+}] + c_3\,(y_0^2 + y_N^2).$$

De la desigualdad (36) del lema (15) del capítulo V para $\varepsilon = 1$ hallaremos, que

$$y_0^2 + y_N^2 \leqslant c_4\,[(y,\ y) + (y_{\overline{x}}^2,\ 1)_{\omega+}],\ c_4 = \frac{8 + l^2}{l\sqrt{16 + l^2}}. \quad (12)$$

Por consiguiente, tenemos

$$(A'(u)\,y,\ y) \leqslant \gamma_2\,[(y,\ y) + (y_{\overline{x}}^2,\ 1)_{\omega+}],$$
$$\gamma_2 = c_2 + c_3 c_4. \quad (13)$$

Definamos en el espacio $H = H(\overline{\omega})$ el operador lineal R, que aplica H sobre H, por las fórmulas

$$Ry_i = \begin{cases} -\dfrac{2}{h}\,y_{\overline{x},\,0} + y_0, & i = 0, \\ -y_{\overline{x}x,\,i} + y_i, & 1 \leqslant i \leqslant N-1, \\ \dfrac{2}{h}\,y_{\overline{x},\,N} + y_N, & i = N. \end{cases}$$

De la primera fórmula de Green de diferencias hallaremos

$$(Ry,\ y) = (y,\ y) + (x_{\overline{x}}^2,\ 1)_{\omega+}. \quad (14)$$

Entonces de (11), (13) y (14) se deduce fácilmente que siendo cumplidas las condiciones (6), (7) para la derivada de Gato $A'(u)$ del operador A son válidas las desigualdades

$$\gamma_1 (Ry, y) \leqslant (A'y, y) \leqslant \gamma_2 (Ry, y), \qquad (15)$$

donde $\gamma_1 = c_1 > 0$ y $\gamma_2 = c_2 + c_3 c_4$, es decir, los operadores R y A' son energéticamente equivalentes con constantes que no dependen del paso h de la red.

Recordemos que más arriba fue obtenida la desigualdad

$$(Au - Av, u - v) \geqslant c_1 [\| u - v \|^2 + ((u - v)_{\bar{x}}^2, 1)_{\omega^+}],$$

si se cumplen las condiciones (2) y (3). De aquí y de (14) se desprende que al cumplirse las condiciones (2) y (3) tiene lugar la estimación

$$(Au - Av, u - v) \geqslant \gamma_1 (R(u - v), u - v), \quad \gamma_1 = c_1 > 0. \qquad (16)$$

Mostremos ahora, que para cualesquiera $u, v \in H$ es cierta la estimación

$$(R^{-1}(Au - Av), Au - Av) \leqslant \gamma_2 (Au - Av, u - v), \qquad (17)$$

donde $\gamma_2 = c_2 (1 + c_4)$, si se cumplen las condiciones

$$\sum_{\alpha=0}^{1} [k_\alpha (x, p_0, p_1) - k_\alpha (x, q_0, q_1)]^2 \leqslant$$

$$\leqslant c_2 \sum_{\alpha=0}^{1} [k_\alpha (x, p_0, p_1) - k_\alpha (x, q_0, q_1)] (p_\alpha - q_\alpha), \qquad (18)$$

$$[\varkappa_\alpha (p_0) - \varkappa_\alpha (q_0)]^2 \leqslant c_2 [\varkappa_\alpha (p_0) - \varkappa_\alpha (q_0)] (p_0 - q_0).$$

Realmente, para la demostración de (17) es suficiente obtener la estimación

$$(Au - Av, z)^2 \leqslant \gamma_2 (Au - Av, u - v)(Rz, z), \qquad (19)$$

para cualesquiera $u, v, z \in H$. Entonces, poniendo aquí $z = R^{-1}(Au - Av)$, tendremos (17). Designemos:

$$p_0 = \bar{p}_0 = u_t, \quad q_0 = \bar{q}_0 = v_t, \quad s_0 = \bar{s}_0 = z_t,$$

$$p_1 = u_{x,\,t}, \quad \bar{p}_1 = u_{\bar{x},\,t}, \quad q_1 = v_{x,\,t}, \quad \bar{q}_1 = v_{\bar{x},\,t},$$

$$s_1 = z_{x,\,t}, \quad \bar{s}_1 = z_{\bar{x},\,t}.$$

Utilizando la definición del operador A y las fórmulas de sumación por partes, obtenemos

$$(Au - Av,\ z)^2 = (\Delta v - \Delta u,\ z)^2 =$$

$$= \left\{ {}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha(x,\ p_0,\ p_1) - k_\alpha(x,\ q_0,\ q_1)],\ s_\alpha)_{\omega-} + \right.$$

$$+ \frac{1}{2} \sum_{\alpha=0}^{1} ([k_\alpha(x,\ \bar{p}_0,\ \bar{p}_1) - k_\alpha(x,\ \bar{q}_0,\ \bar{q}_1)],\ \bar{s}_\alpha)_{\omega+} +$$

$$\left. + [\varkappa_1(\bar{p}_0,) - \varkappa_1(\bar{q}_0)]\,\bar{s}_0\,|_{i=N} + [\varkappa_0(p_0) - \varkappa_0(q_0)]\,s_0\,|_{i=0} \right\}^2 .$$

Utilizando la desigualdad de Cauchy-Buniakovsky hallaremos sucesivamente

$$(Au - Av,\ z)^2 \leqslant \left[{}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha'(x,\ p_0,\ p_1) - \right.$$

$$- k_\alpha(x,\ q_0,\ q_1)]^2,\ 1)_{\omega-}^{1/2} (s_\alpha^2,\ 1)_{\omega-}^{1/2} +$$

$$+ {}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha(x,\ \bar{p}_0,\ \bar{p}_1) - k_\alpha(x,\ \bar{q}_0,\ \bar{q}_1)]^2,\ 1)_{\omega+}^{1/2} \times$$

$$\times (\bar{s}_\alpha^2,\ 1)_{\omega+}^{1/2} + [\varkappa_1(\bar{p}_0) - \varkappa_1(\bar{q}_0)]\bar{s}_0\,|_{i=N} +$$

$$\left. + [\varkappa_0(p_0) - \varkappa_0(q_0)]\,s_0\,|_{i=0} \right|^2 \leqslant$$

$$\leqslant \{ {}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha(x,\ p_0,\ p_1) - k_\alpha(x,\ q_0,\ q_1)]^2,\ 1)_{\omega-} +$$

$$+ {}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha(x,\ \bar{p}_0,\ \bar{p}_1) - k_\alpha(x,\ \bar{q}_0,\ \bar{q}_1)]^2,\ 1)_{\omega+} +$$

$$+ [\varkappa_1(\bar{p}_0) - \varkappa_1(\bar{q}_0)]^2_{i=N} + [\varkappa_0(p_0) -$$

$$- \varkappa_0(q_0)]^2_{i=0} \} \{ {}^1/_2 \sum_{\alpha=0}^{1} [s_\alpha^2,\ 1)_{\omega-} + (\bar{s}_\alpha^2,\ 1)_{\omega+}] + \bar{s}_0^2\,|_{i=N} + s_0^2\,|_{i=0} \}.$$

Teniendo en cuenta, que es cierta la igualdad

$$(Au - Av,\ u - v) = {}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha(x,\ p_0,\ p_1) -$$

$$- k_\alpha(x,\ q_0,\ q_1,\ )](p_\alpha - q_\alpha),\ 1)_{\omega-} +$$

$$+ {}^1/_2 \sum_{\alpha=0}^{1} ([k_\alpha(x,\ \bar{p}_0,\ \bar{p}_1) - k_\alpha(x,\ \bar{q}_0,\ \bar{q}_1)] \times$$

$$\times (\bar{p}_\alpha - \bar{q}_\alpha),\ 1)_{\omega+} + [\varkappa_1(\bar{p}_0) - \varkappa_1(\bar{q}_0)] \times$$

$$\times (\bar{p}_0 - \bar{q}_0)|_{i=N} + [\varkappa_0(p_0) - \varkappa_0(q_0)](p_0 - q_0)|_{i=0},$$

y también que en virtud de (12), (14) y de las notaciones introducidas

$${}^1/_2 \sum_{\alpha=0}^{1} [(s_\alpha^2,\ 1)_{\omega-} + (\bar{s}_\alpha^2, 1)_{\omega+}] + \bar{s}_0^2|_{i=N} + s_0^2|_{i=0} =$$

$$= {}^1/_2[(z^2,\ 1)_{\omega+} + (z^2,\ 1)_\omega + (z_x^2,\ 1)_{\omega+} +$$

$$+ (z_x^2,\ 1)_{\omega-}] + z_N^2 + z_0^2 = (z^2,\ 1) + z_x^2,\ 1)_{\omega+} +$$

$$+ z_N^2 + z_0^2 \leqslant (1 + c_4)(Rz,\ z),$$

obtendremos la estimación (19), si se cumplen las condiciones (18). La afirmación está demostrada.

2. Método de iteración simple. Examinemos ahora los métodos iterativos de resolución del esquema de diferencias no lineal (4) construido. Supongamos primeramente que se cumplen las condiciones (2), (3) y (18).

Para resolver las ecuaciones (4) utilizaremos el método implícito de iteración simple

$$B\frac{y_{k+1} - y_k}{\tau} + Ay_k = f,\ k = 0,\ 1,\ \ldots,\ y_0 \in H, \qquad (20)$$

donde $A = -\Lambda$, $B = R$ y el operador R está definido más arriba. De (20) se sigue, que para encontrar y_{k+1} siendo prefijada y_k se exige resolver la ecuación lineal

$$By_{k+1} = \varphi,\quad \varphi = By_k - \tau(Ay_k - f)$$

ó en forma detallada

$$-y_{k+1}(i-1) + cy_{k+1}(i) - y_{k+1}(i+1) = h^2\varphi(i),$$

$$1 \leqslant i \leqslant N - 1,$$

$$cy_{k+1}(0) - 2y_{k+1}(1) = h^2\varphi(0),\quad i = 0,$$

$$-2y_{k+1}(N-1) + cy_{k+1}(N) = h^2\varphi(N),\quad i = N,$$

donde $c = 2 + h^2$. Ya que $c > 2$, entonces el problema de contorno de diferencias puede ser resuelto por el método de factorización monótona con un gasto de $O(N)$ operaciones aritméticas.

Queda por indicar el valor del parámetro iterativo τ y estimar el número de iteraciones exigidas. Puesto que son cumplidas las condiciones (2), (3) y (18), tienen lugar las estimaciones (16) y (17), las cuales se pueden escribir en la forma

$$(Au - Av,\ u - v) \geqslant \gamma_1 (B(u - v),\ u - v),\quad \gamma_1 = c_1 > 0,$$
$$(B^{-1}(Au - Av),\ Au - Av) \leqslant \gamma_2 (Au - Av,\ u - v),$$
$$\gamma_2 = c_2 (1 + c_4), \qquad (21)$$

donde c_1 está dado en (2), c_2 en (18) y c_4 en (12).

Como el operador B es autoconjugado y definido positivo, entonces investigaremos la convergencia del método (20) en el espacio energético H_D, donde $D = B$. Para la elección indicada del operador D las desigualdades (21) coinciden con las desigualdades (4) y (5). Por eso para elegir el parámetro iterativo τ es posible utilizar el teorema 1. Obtendremos que para $\tau = 1/\gamma_2 = 1/(c_2(1 + c_4))$ el método iterativo (20) converge en H_D, y para el error es cierta la estimación $\| y_n - u \|_B \leqslant \rho^n \| y_0 - u \|_B$, $\rho = \sqrt{1 - \xi}$, $\xi = \gamma_1/\gamma_2$ con cualquier aproximación inicial y_0.

Así pues, si se cumplen las condiciones (2), (3), (18), entonces el método iterativo de iteración simple (20) con el valor indicado del parámetro τ permite obtener la solución del esquema de diferencias (4) con una exactitud ε por $n \geqslant n_0(\varepsilon)$ iteraciones, donde

$$n_0(\varepsilon) = \frac{\ln \varepsilon}{\ln \rho} = \frac{2 \ln \varepsilon}{\ln \left(1 - \frac{c_1}{c_2 (1 + c_4)}\right)}.$$

Ya que las constantes c_1, c_2 y c_4 no dependen del paso h de la red, entonces el número de iteraciones $n_0(\varepsilon)$ depende sólo de ε y no cambia con la reducción de la red.

Examinemos ahora el método iterativo (20) bajo la suposición de que se cumplen (6) y (7) para las derivadas $a_{\alpha\beta} = \partial k_\alpha / \partial p_\beta$ y $\sigma_\alpha = \partial \varkappa_\alpha / \partial p_0$, y también las condiciones de simetría (9). Entonces para la derivada de Gato del operador A serán válidas las desigualdades (15), las cuales en virtud de la elección $B = R$ se pueden escribir en la forma

$$\gamma_1 (By,\ y) \leqslant (A'(v)\, y,\ y) \leqslant \gamma_2 (By,\ y), \quad v,\ y \in H, \qquad (22)$$

donde $\gamma_1 = c_1$, $\gamma_2 = c_2 + c_3 c_4$, c_1, c_2 y c_3 están definidas en (6), (7), y c_4 en (12).

Sea $D = B$. Entonces el operador $DB^{-1}A'(v) = A'(v)$ en virtud del lema 4 será autoconjugado en H, y, por consiguiente, se cumplen las condiciones del teorema 2, y las desigualdades (22) coinciden con las desigualdades (14). Por lo tanto el parámetro τ en el esquema (20) se debe tomar igual a $\tau = \tau_0 = 2/(\gamma_1 + \gamma_2)$. Con esto para el error $y_n - u$ y para el número de iteraciones serán válidas las estimaciones

$$\|y_n - u\|_B \leqslant \rho_0^n \|y_0 - u\|, \quad \rho_0 = \frac{1-\xi}{1+\xi},$$

$$\xi = \frac{\gamma_1}{\gamma_2} = \frac{c_1}{c_2 + c_3 c_4}, \quad n \geqslant n_0(\varepsilon) = \ln \varepsilon / \ln \rho_0.$$

Aquí, al igual que para el método interior, el número de iteraciones no depende del paso h de la red. Para el operador B elegido en virtud de la primera fórmula de Green de diferencias tendremos las siguientes representaciones para la norma $\|z\|_B$:

$$\|z\|_B^2 = (z, z) + (z_{\bar{x}}^2, 1)_{\omega^+}.$$

Nosotros examinamos los métodos de resolución del esquema de diferencias no lineal que aproxima una ecuación cuasilineal unidimensional sobre una red uniforme. No representa trabajo transferir estos exámenes al caso de una red no uniforme arbitraria, y también a los esquemas de diferencias que aproximan los problemas de contorno fundamentales para una ecuación cuasilineal elíptica de segundo orden en un rectángulo.

3. Métodos iterativos para ecuaciones cuasilineales elípticas en diferencias en un rectángulo. Para el rectángulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \alpha = 1, 2\}$ con frontera Γ se exige hallar la solución de la ecuación

$$\sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha} k_\alpha \left(x, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right) -$$

$$- k_0 \left(x, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right) = -\varphi(x), \quad x \in G, \quad (23)$$

que satisfaga las condiciones de contorno de tercer género

$$k_\alpha\left(x,\ u,\ \frac{\partial u}{\partial x_1},\ \frac{\partial u}{\partial x_2}\right)=\varkappa_{-\alpha}(x,\ u)-g_{-\alpha}(x),\ x_\alpha=0.$$

$$-k_\alpha\left(x,\ u,\ \frac{\partial u}{\partial x_1},\ \frac{\partial u}{\partial x_2}\right)= \tag{24}$$

$$=\varkappa_{+\alpha}(x,\ u)-g_{+\alpha}(x),\ x_\alpha=l_\alpha,\quad \alpha=1,\ 2.$$

Supongamos, como en el caso unidimensional, que son cumplidas las condiciones siguientes. Las funciones $k_\alpha(x,\ p)$ y $\varkappa_{\pm\alpha}(x,\ p_0)$ son continuas respecto de $p=(p_0,\ p_1,\ p_2)$ y p_0 y, además

$$\sum_{\alpha=0}^{2}[k_\alpha(x,\ p)-k_\alpha(x,\ q)](p_\alpha-q_\alpha)\leqslant c_1\sum_{\alpha=0}^{2}(p_\alpha-q_\alpha)^2,$$

$$\sum_{\alpha=0}^{2}[k_\alpha(x,\ p)-k_\alpha(x,\ q)]^2\leqslant$$

$$\leqslant c_2\sum_{\alpha=0}^{2}[k_\alpha(x,\ p)-k_\alpha(x,\ q)](p_\alpha-q_\alpha),$$

$$[\varkappa_{\pm\alpha}(x,\ p_0)-\varkappa_{\pm\alpha}(q_0)]^2\leqslant c_2[\varkappa_{\pm\alpha}(x,\ p_0)-\varkappa_{\pm\alpha}(q_0)]\times$$

$$\times(p_0-q_0),\ \alpha=1,\ 2,$$

donde $c_1>0$ y $c_2>0$, $x\in\overline{G}$ y $|p|,\ |q|<\infty$.

Introduzcamos en la región $\overline{G}$ la red rectangular uniforme

$$\overline{\omega}=\{x_{ij}=(ih_1,\ jh_2),\quad 0\leqslant i\leqslant N_1,\quad 0\leqslant j\leqslant N_2,$$

$$h_\alpha N_\alpha=l_\alpha,\quad \alpha=1,\ 2\}.$$

El esquema de diferencias más simple, correspondiente al problema (23), (24), tiene la forma

$$\Lambda y=-f,\quad x\in\overline{\omega}, \tag{25}$$

$$\Lambda=\Lambda_1+\Lambda_2,\quad f=\varphi+2\varphi_1/h_1+2\varphi_2/h_2,$$

donde

$$\varphi_\alpha(x)=\begin{cases}g_{-\alpha}(x), & x_\alpha=0,\\ 0, & h_\alpha\leqslant x_\alpha\leqslant l_\alpha-h_\alpha,\\ g_{+\alpha}(x), & x_\alpha=l_\alpha,\quad 0\leqslant x_{3-\alpha}\leqslant l_{3-\alpha},\end{cases}$$

y los operadores Λ_α, $\alpha=1,\ 2$, están definidos mediante las fórmulas:

1) para $h_\beta \leqslant x_\beta \leqslant l_\beta - h_\beta$ tenemos

$$\Lambda_\alpha y = 1/2\{[k_\alpha(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2})]_{x_\alpha} + [k_\alpha(x,\ y,\ y_{x_1},\ y_{x_2})]_{\bar{x}_\alpha}\} - \\ - 1/4\,[k_0(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) + k_0(x,\ y,\ y_{x_1},\ y_{x_2})],$$

$h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha$;

$$\Lambda_\alpha y = \frac{1}{h_\alpha}[k_\alpha^{+1\alpha}(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) + k_\alpha(x,\ y,\ y_{x_1},\ y_{x_2})] - \\ - 1/2 k_0(x,\ y,\ y_{x_1},\ y_{x_2}) - \frac{2}{h_\alpha}\varkappa_{-\alpha}(x,\ y),\ x_\alpha = 0;$$

$$\Lambda_\alpha y = -\frac{1}{h_\alpha}[k_\alpha(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) + k_\alpha^{-1\alpha}(x,\ y,\ y_{x_1},\ y_{x_2})] - \\ - 1/2\,(k_0(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) - \frac{2}{h_\alpha}\varkappa_{+\alpha}(x,\ y),\quad x_\alpha = l_\alpha;$$

2) para $x_\beta = 0$ tenemos

$$\Lambda_\alpha y = [k_\alpha(x,\ y,\ y_{x_1},\ y_{x_2})]_{\bar{x}_\alpha} - 1/2 k_0(x,\ y,\ y_{x_1},\ y_{x_2})$$

siendo $h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha$;

$$\Lambda_\alpha y = \frac{2}{h_\alpha}k_\alpha(x,\ y,\ y_{x_1},\ y_{x_2}) - k_0(x,\ y,\ y_{x_1},\ y_{x_2}) - \\ - \frac{2}{h_\alpha}\varkappa_{-\alpha}(x,\ y)$$

siendo $x_\alpha = 0$;

$$\Lambda_\alpha y = -\frac{2}{h_\alpha}k_\alpha^{1\alpha}(x,\ y,\ y_{x_1},\ y_{x_2}) - \frac{2}{h_\alpha}\varkappa_{+\alpha}(x,\ y),\quad x_\alpha = l_\alpha;$$

3) para $x_\beta = l_\beta$ tenemos

$$\Lambda_\alpha y = [k_\alpha(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2})]_{x_\alpha} - 1/2 k_0(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2})$$

siendo $h_\alpha \leqslant x_\alpha \leqslant l_\alpha - h_\alpha$;

$$\Lambda_\alpha y = \frac{2}{h_\alpha}k_\alpha^{+1\alpha}(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) - \frac{2}{h_\alpha}\varkappa_{-\alpha}(x,\ y),\ x_\alpha = 0;$$

$$\Lambda_\alpha y = -\frac{2}{h_\alpha}k_\alpha(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) - \\ - k_0(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2}) - \frac{2}{h_\alpha}\varkappa_{+\alpha}(x,\ y)$$

siendo $x_\alpha = l_\alpha$.

Aquí $\beta = 3 - \alpha$, $\alpha = 1, 2$ y son utilizadas las notaciones

$$k_1^{+1_1}(x,\ y,\ y_{\bar{x}_1},\ y_{\bar{x}_2})\,|\,x_{ij} = \\ = k_1(x_{i+1},\ j,\ y(i+1,\ j),\ y_{\bar{x}_1}(i+1,\ j),\ y_{\bar{x}_2}(i+1,\ j),$$

y también notaciones análogas para $k_1^{-1_1}$ y $k_2^{\pm 1_2}$.

En el espacio H de las funciones reticulares definidas sobre $\overline{\omega}$, determinaremos el producto escalar

$$(u,\ v)=\sum_{i=0}^{N_1}\sum_{j=0}^{N_2}\hbar_1(i)\,\hbar_2(j)\,u(i,\ j)\,v(i,\ j),$$

$$\hbar_\alpha(k)=\begin{cases}h_\alpha, & 1\leqslant k\leqslant N_\alpha-1,\\ 0,\ 5h_\alpha, & k=0,\ N_\alpha\end{cases}$$

y los operadores $A_\alpha=-\Lambda_\alpha$, $\alpha=1,\ 2$, $A=A_1+A_2$, $R=R_1+R_2$, donde

$$R_\alpha y=\begin{cases}-\dfrac{2}{h_\alpha}y_{x_\alpha}+{}^1/_2y, & x_\alpha=0,\\ -y_{\overline{x}_\alpha x_\alpha}+{}^1/_2y, & h_\alpha\leqslant x_\alpha\leqslant l_\alpha-h_\alpha,\\ \dfrac{2}{h_\alpha}y_{\overline{x}_\alpha}+{}^1/_2y, & x_\alpha=l_\alpha,\quad \alpha=1,\ 2\end{cases}$$

y $0\leqslant x_\beta\leqslant l_\beta$. Entonces el esquema de diferencias (25) se escribe en la forma de la ecuación operacional

$$Au=f \tag{26}$$

con el operador no lineal A.

Utilizando las suposiciones hechas más arriba con respecto a los coeficientes $k_\alpha(x,\ p)$ y $\varkappa_{\pm\alpha}(p_0)$, al igual que en el caso unidimensional, obtendremos que tienen lugar las desigualdades (16) y (17), donde c_4 es la constante de la desigualdad

$$\begin{aligned}\sum_{j=0}^{N_2}\hbar_2(j)\,[y^2(0,\ j)+y^2(N_1,\ j)]+&\\ +\sum_{i=0}^{N_1}\hbar_1(i)\,[y^2(i,\ 0)+y^2(i,\ N_2)]\leqslant&\\ \leqslant c_4\Big[\sum_{i=0}^{N_1}\sum_{j=0}^{N_2}y^2(i,\ j)\,\hbar_1(i)\,\hbar_2(j)+&\\ +\sum_{j=0}^{N_2}\sum_{i=1}^{N_1}h_1\hbar_2(j)\,y_{x_1}^2(i,\ j)+&\\ +\sum_{i=0}^{N_1}\sum_{j=1}^{N_2}\hbar_1(i)\,h_2y_{x_2}^2(i,\ j)\Big].&\end{aligned} \tag{27}$$

Mostremos, que

$$c_4=\sqrt{2}\,(16+l^2)/\left(l\sqrt{32+l^2}\right),\quad l=\min(l_1,\ l_2). \tag{28}$$

En efecto, de la desigualdad (36) del lema 15 del capítulo V para $\varepsilon = \sqrt{2}$ obtenemos

$$y^2(0,\ j) + y^2(N_1,\ j) \leqslant \frac{(16 + l_1^2)\sqrt{2}}{l_1\sqrt{32 + l_1^2}} \times$$
$$\times \Big[\sum_{i=1}^{N_1} h_1 y_{\bar{x}_1}^2 (i,\ j) + {}^1/_2 \sum_{i=0}^{N_1} \hbar_1 (i)\, y^2 (i,\ j) \Big].$$

Señalemos que si aquí cambiamos l_1 por l, entonces la desigualdad sólo se fortalece. Multipliquemos ahora los primero y segundo miembros de la desigualdad obtenida por $\hbar_2 (j)$ y sumemos respecto de j desde 0 hasta N_2. Tendremos

$$\sum_{j=0}^{N_2} \hbar_2 (j) [y^2 (0,\ j) + y^2 (N_1,\ j)] \leqslant$$
$$\leqslant c_4 \Big[\sum_{j=0}^{N_2} \sum_{i=0}^{N_1} h_1 \hbar_2 (j)\, y_{\bar{x}_1}^2 (i,\ j) +$$
$$+ {}^1/_2 \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} y^2 (i,\ j)\, \hbar_1 (i)\, \hbar_2 (j) \Big], \qquad (29)$$

donde c_4 está definida en (28). Análogamente hallaremos

$$\sum_{i=0}^{N_1} \hbar_1 (i) [y^2 (i,\ 0) + y^2 (i,\ N_2)] \leqslant$$
$$\leqslant c_4 \Big[\sum_{i=0}^{N_1} \sum_{j=1}^{N_2} \hbar_1 (i)\, h_2 y_{\bar{x}_2}^2 (i,\ j) +$$
$$+ {}^1/_2 \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} y^2 (i,\ j)\, \hbar_1 (i)\, \hbar_2 (j) \Big]. \qquad (30)$$

Adicionando (29) y (30), obtendremos la desigualdad (27).

Para resolver la ecuación (26) se puede aprovechar el método implícito de iteración simple (20), donde $B = R$ y $\tau = 1/\gamma_2 = 1/(c_2\,(1 + c_4))$. Entonces en virtud del teorema 1 el método iterativo (20) convergerá en H_B y para el error será cierta la estimación

$$\| y_n - u \|_B \leqslant \rho^n \| y_0 - u \|_B, \quad \rho = \sqrt{1 - \xi},$$
$$\xi = \gamma_1/\gamma_2 = c_1/(c_2\,(1 + c_4)).$$

Por lo tanto, el número de iteraciones $n_0 (\varepsilon)$, que se deben cumplir para alcanzar una exactitud relativa ε, no dependerá del número de nodos de la red $\overline{\omega}$.

Para encontrar y_{k+1} tenemos el problema

$$Ry_{k+1} = \varphi, \quad \varphi = Ry_k - \tau (Ay_k - f).$$

Ya que el operador R corresponde al segundo problema de contorno para una ecuación en diferencias con coeficientes constantes, entonces el problema indicado puede ser resuelto por los métodos directos, descritos en los capítulos III y IV, con un gasto de $O(N^2 \log_2 N)$ operaciones aritméticas ($N_1 = N_2 = N = 2^n$). Si las funciones $k_\alpha(x, p)$ y $\varkappa_{\pm\alpha}(x, p_0)$ son diferenciables, entonces el operador A tiene derivada de Gato, la cual es un operador autoconjugado en H, si se cumplen las condiciones

$$a_{\alpha\beta}(x, p) = a_{\beta\alpha}(x, p), \quad \alpha, \beta = 0, 1, 2, \tag{31}$$

donde $a_{\alpha\beta}(x, p) = \frac{\partial k_\alpha(x, p)}{\partial p_\beta}$. Se puede mostrar, que si además de (31) se cumplen las condiciones

$$c_1 \sum_{\alpha=0}^{2} \xi_\alpha^2 \leqslant \sum_{\alpha,\beta=0}^{2} a_{\alpha\beta}(x, p)\, \xi_\alpha \xi_\beta \leqslant c_2 \sum_{\alpha=0}^{2} \xi_\alpha^2, \quad c_1 > 0,$$

$$0 \leqslant \frac{\partial \varkappa_{\pm\alpha}(x, p_0)}{\partial p_0} \leqslant c_3, \quad \alpha = 1, 2,$$

entonces son ciertas las desigualdades (15), donde $\gamma_1 = c_1$, $\gamma_2 = c_2 + c_3 c_4$, mientras que c_4 está definido en (28). Entonces en el método iterativo (20) con $B = R$ el parámetro τ se puede tomar igual a $\tau_0 = 2/(\gamma_1 + \gamma_2)$. En virtud del teorema 4 para el error será cierta la estimación

$$\| y_n - u \|_B \leqslant \rho_0^n \| y_0 - u \|_B, \quad \rho_0 = (1 - \xi)/(1 + \xi),$$

$$\xi = \gamma_1/\gamma_2.$$

Supongamos que ahora se exige hallar la solución del primer problema de contorno en un rectángulo $\overline{G}$

$$\sum_{\alpha=1}^{2} \frac{\partial}{\partial x_\alpha} k_\alpha\left(x, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right) -$$

$$- k_0\left(x, u, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}\right) = -\varphi(x), \quad x \in G, \tag{32}$$

$$u(x) = 0, \quad x \in \Gamma.$$

Supongamos que las funciones $k_\alpha(x, p)$ son continuas respecto a $p = (p_0, p_1, p_2)$ y que se cumplen las condiciones

$$\sum_{\alpha=1}^{2} [k_\alpha(x, p) - k_\alpha(x, q)](p_\alpha - q_\alpha) \geqslant c_1 \sum_{\alpha=1}^{2} (p_\alpha - q_\alpha)^2,$$

$$[k_0(x, p) - k_0(x, q)](p_0 - q_0) \geqslant 0, \tag{33}$$

$$\sum_{\alpha=0}^{2} [k_\alpha(x, p) - k_\alpha(x, q)]^2 \leqslant c_2 \sum_{\alpha=0}^{2} [k_\alpha(x, p) - k_\alpha(x, q)] \times (p_\alpha - q_\alpha),$$

donde $c_1 > 0$, $c_2 > 0$ para $x \in \overline{G}$ y $|p|, |q| < \infty$.

Al problema (32) sobre la red rectangular uniforme $\overline{\omega} = \omega \cup \gamma$, introducida anteriormente, le pondremos en correspondencia el esquema de diferencias

$$\Lambda y = -f, \ x \in \omega, \ y(x) = 0, \quad x \in \gamma, \tag{34}$$

donde $f = \varphi$, y el operador en diferencias Λ está definido de la forma siguiente:

$$\begin{aligned}\Lambda y = \Lambda - y = {}^1/_2 \{ & [k_1(x, y, y_{\overline{x}_1}, y_{\overline{x}_2})]_{x_1} + \\ & + [k_1(x, y, y_{x_1}, y_{x_2})]_{\overline{x}_1} + [k_2(x, y, y_{\overline{x}_1}, y_{\overline{x}_2})]_{x_2} + \\ & + [k_2(x, y, y_{x_1}, y_{x_2})]_{\overline{x}_2} - k_0(x, y, y_{\overline{x}_1}, y_{\overline{x}_2}) - \\ & - k_0(x, y, y_{x_1}, y_{x_2}) \}.\end{aligned}$$

Mostremos otras dos posibles aproximaciones:

$$\begin{aligned}\Lambda y = \Lambda^+ y = {}^1/_2 \{ & [k_1(x, y, y_{\overline{x}_1}, y_{x_2})]_{x_1} + \\ & + [k_1(x, y, y_{x_1}, y_{\overline{x}_2})]_{\overline{x}_1} + [k_2(x, y, y_{x_1}, y_{\overline{x}_2})]_{x_2} + \\ & + [k_2(x, y, y_{\overline{x}_1}, y_{x_2})]_{\overline{x}_2} - k_0(x, y, y_{\overline{x}_1}, y_{x_2}) - \\ & - k_0(x, y, y_{x_1}, y_{\overline{x}_2}) \}\end{aligned}$$

y $\Lambda = {}^1/_2 (\Lambda^- + \Lambda^+)$.

Para el ejemplo dado H es el espacio de las funciones reticulares definidas sobre ω, en el cual el producto escalar está definido por la fórmula

$$(u, v) = \sum_{i=1}^{N_1-1} \sum_{j=1}^{N_2-1} h_1 h_2 u(i, j) v(i, j).$$

Si en las ecuaciones del esquema (34) ponemos $y|_\gamma = 0$, obtendremos el esquema de diferencias $\overline{\Lambda} y = -f$. Definiendo

el operador $A = -\overline{\Lambda}$, escribamos el esquema obtenido en la forma de la ecuación operacional (26) en el espacio H.

Utilizando las condiciones (33), obtendremos para todas las tres aproximaciones que el operador A satisface las desigualdades (16) y (17):

$$(Au - Av,\ u - v) \geqslant \gamma_1 (R(u - v),\ u - v),\quad \gamma_1 = c_1 > 0$$

$$(R^{-1}(Au - Av),\ Au - Av) \leqslant \gamma_2 (Au - Av,\ u - v),$$

$$\gamma_2 = c_2 (1 + c_4),$$

donde

$$c_4 = \frac{1}{\delta},\ \delta = \frac{4}{h_1^2} \operatorname{sen}^2 \frac{\pi h_1}{2l_1} + \frac{4}{h_2^2} \operatorname{sen}^2 \frac{\pi h_2}{2l_2} \geqslant \frac{8}{l_1^2} + \frac{8}{l_2^2},$$

y el operador R corresponde al operador de Laplace en diferencias $Ry = -\mathring{\mathcal{R}}\mathring{y}$, $y(x) = \mathring{y}(x)$ para $x \in w$ y $\mathring{y}(x) = 0$ para $x \in \gamma$, $\mathcal{R}u = u_{\bar{x}_1 x_1} + u_{\bar{x}_2 x_2}$.

Para resolver la ecuación (26) utilicemos el método de iteración simple (20) con $B = R$ y $\tau = 1/\gamma_2$. En virtud del teorema 1 tendremos la estimación

$$\| y_n - u \|_B \leqslant \rho^n \| y_0 - u \|_B,\quad \rho = \sqrt{1 - \xi},\ \xi = \gamma_1/\gamma_2.$$

Como antes, para resolver la ecuación $Ry_{k+1} = Ry_k - \tau (Ay_k - f)$ se pueden utilizar los métodos directos de reducción completa o de separación de variables, propuestos en los capítulos III y IV.

4. Métodos iterativos para ecuaciones débilmente no lineales. En el rectángulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$ examinaremos la ecuación elíptica de segundo orden débilmente no lineal

$$Lu = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} - k_0\left(x,\ u,\ \frac{\partial u}{\partial x_1},\ \frac{\partial u}{\partial x_2}\right) = 0,\quad x \in G \qquad (35)$$

con las condiciones de contorno de primer género

$$u(x) = 0,\quad x \in \Gamma. \qquad (36)$$

La *no linealidad débil* de la ecuación (35) significa que la función $k_0 (x, p_0, p_1, p_2)$ está definida para $x \in G$ y $|p_0|$, $|p_1|$, $|p_2| < \infty$, es continua respecto a x para p_0, p_1 y p_2 fijos, y también existen las derivadas de la función $k_0 (x, p_0, p_1, p_2)$ por p_0, p_1 y p_2, las cuales satisfacen las condiciones

$$c_2 \geqslant \frac{\partial k_0}{\partial p_0} \geqslant 0,\quad \left|\frac{\partial k_0}{\partial p_\alpha}\right| \leqslant M,\quad \alpha = 1, 2. \qquad (37)$$

Sobre la red rectangular uniforme $\overline{\omega} = \omega \cup \gamma$, introducida anteriormente, el esquema de diferencias correspondiente al problema (35), (36) tiene la forma

$$\Lambda y = 0, \quad x \in \omega, \quad y(x) = 0, \quad x \in \gamma, \tag{38}$$
$$\Lambda y = \mathcal{R} y - {}^1/_2 [k_0(x, y, y_{\bar{x}_1}, y_{\bar{x}_2}) + k_0(x, y, y_{x_1}, y_{x_2})],$$

donde $\mathcal{R} y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2}$ es el operador de Laplace de diferencias.

Definamos ahora el operador de diferencias $\Lambda'(v)$, dependiente de v:

$$\Lambda'(v)\, y = \mathcal{R} y - {}^1/_2 [a_{01}(x, v, v_{\bar{x}_1}, v_{\bar{x}_2})\, y_{\bar{x}_1} +$$
$$+ a_{01}(x, v, v_{x_1}, v_{x_2})\, y_{x_1} + a_{02}(x, v, v_{\bar{x}_1}, v_{\bar{x}_2})\, y_{\bar{x}_2} +$$
$$+ a_{02}(x, v, v_{x_1}, v_{x_2})\, y_{x_2} + (a_{00}(x, v, v_{\bar{x}_1}, v_{\bar{x}_2}) +$$
$$+ a_{00}(x, v, v_{x_1}, v_{x_2}))\, y],$$

donde

$$a_{0\alpha}(x, p_0, p_1, p_2) = \frac{\partial k_0(x, p_0, p_1, p_2)}{\partial p_\alpha}, \quad \alpha = 0, 1, 2.$$

En el espacio H de las funciones reticulares prefijadas sobre ω definiremos los operadores:

$$Ay = -\Lambda \mathring{y}, \quad Ry = -\mathcal{R} \mathring{y}, \quad A'(v)\, y = -\Lambda'(\mathring{v})\, \mathring{y},$$

donde

$$y(x) = \mathring{y}(x), \quad v(x) = \mathring{v}(x) \quad \text{para} \quad x \in \omega$$

e

$$\mathring{y}(x) = 0, \quad \mathring{v}(x) = 0 \quad \text{para} \quad x \in \gamma.$$

El operador $A'(v)$ es la derivada de Gato del operador A. Utilizando estas notaciones, escribamos el esquema de diferencias en la forma de la ecuación operacional (26).

Si $k_0(x, p_0, p_1, p_2)$ no depende de p_1 y p_2, es decir

$$k_0(x, p_0, p_1, p_2) = k_0(x, p_0),$$

entonces

$$a_{01}(x, p) = a_{02}(x, p) = 0.$$

En este caso el operador $A'(v)$ es autoconjugado en H.

Utilizando la estimación por debajo para el operador de diferencias $(-\mathscr{R})$

$$(-\mathscr{R}\overset{\circ}{y},\ \overset{\circ}{y}) = -(\overset{\circ}{y}_{\bar{x}_1 x_1} + \overset{\circ}{y}_{\bar{x}_2 x_2},\ \overset{\circ}{y}) \geqslant \delta\,(\overset{\circ}{y},\ \overset{\circ}{y}),$$

donde

$$\delta = \frac{4}{h_1^2}\operatorname{sen}^2\frac{\pi h_1}{2l_1} + \frac{4}{h_1^2}\operatorname{sen}^2\frac{\pi h_2}{2l_2} \geqslant \frac{8}{l_1^2} + \frac{8}{l_2^2},$$

las condiciones (37) para $M = 0$ y las igualdades

$$-(\Lambda'(\overset{\circ}{v})\,\overset{\circ}{y},\ \overset{\circ}{y}) = -(\mathscr{R}\overset{\circ}{y},\ \overset{\circ}{y}) + (a_{00}(x,\ \overset{\circ}{v})\,\overset{\circ}{y},\ \overset{\circ}{y}),$$

obtendremos

$$\gamma_1\,(Ry,\ y) \leqslant (A'\,(v)\,y,\ y) \leqslant \gamma_2\,(Ry,\ y),$$

donde

$$\gamma_1 = 1, \quad \gamma_2 = 1 + c_2/\delta.$$

Por consiguiente, si para el caso «autoconjugado» examinado se utiliza el método iterativo (20) con $B = D = R$ y $\tau = \tau_0 = 2/(\gamma_1 + \gamma_2)$, entonces en virtud del teorema 2 para el error será cierta la estimación

$$\| y_n - u \|_B \leqslant \rho_0^n \| y_0 - u \|_B,$$
$$\rho_0 = (1 - \xi)/(1 + \xi), \quad \xi = \gamma_1/\gamma_2.$$

El operador R en el esquema (20) se puede invertir por uno de los métodos directos.

Capítulo XIV

Ejemplos de resolución de ecuaciones elípticas reticulares

En el § 1 se examinan algunos procedimientos para construir esquemas iterativos implícitos, en particular, basados en la separación de un regulador. En el § 2 se exponen los métodos directos e iterativos para resolver los problemas de diferencias en el caso multidimensional. Se examina el método iterativo de resolución del tercer problema de contorno para una ecuación elíptica con derivadas mixtas sobre el rectángulo. El § 3 está dedicado a los métodos de resolución de sistemas de ecuaciones elípticas. Los métodos encaminados a resolver el problema de Dirichlet de diferencias para la ecuación de Poisson en la región de forma compuesta, los cuales se basan en reducir el problema inicial a una serie de problemas resolubles en un rectángulo, vienen trazados en el § 4.

§ 1. Métodos de construcción de esquemas iterativos implícitos

1. Principio de regularización en la teoría general de los métodos iterativos. En los capítulos VI—VIII, XII y XIII fue expuesta la teoría general de los métodos iterativos, utilizados para la resolución de la ecuación operacional

$$Au = f. \tag{1}$$

En la teoría general de los métodos iterativos nosotros no utilizamos la estructura concreta de los operadores del esquema iterativo, o sea, la teoría utiliza el mínimo de información de carácter funcional general con respecto a los operadores. Esto permite indicar (para los operadores fijos del esquema) los principios generales de construcción de los métodos iterativos optimales. Por ejemplo, si los operadores

A y B del esquema iterativo de dos capas

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f, \quad k=0, 1, \ldots, y_0\in H \tag{2}$$

satisfacen las condiciones

$$B=B^*>0, \quad A=A^*>0, \tag{3}$$

$$\gamma_1 B\leqslant A\leqslant\gamma_2 B, \quad \gamma_1>0, \tag{4}$$

entonces el conjunto de los parámetros iterativos de Chebishev τ_k:

$$\tau_k=\frac{\tau_0}{1+\rho_0\mu_k}, \quad \mu_k\in\mathfrak{M}_n=$$

$$=\left\{-\cos\frac{(2i-1)\pi}{2n}, \quad 1\leqslant i\leqslant n\right\}, \quad 1\leqslant k\leqslant n,$$

donde

$$\tau_0=\frac{2}{\gamma_1+\gamma_2}, \quad \rho_0=\frac{1-\xi}{1+\xi}, \quad \xi=\frac{\gamma_1}{\gamma_2}$$

es el mejor.

Por cuáles instrucciones debemos guiarnos al elegir el operador B? En el § 3 del capítulo V se señaló, que la elección de B debe ser subordinada a dos requisitos: 1) al aseguramiento de la más rápida convergencia del método; 2) al rendimiento económico en la inversión de este operador.

Para el ejemplo dado más arriba la primera exigencia se cumple, si la energía del operador B será proxima a la energía del operador A, es decir, si en las desigualdades (4) γ_1 y γ_2 serán cercanas. Para satisfacer la segunda exigencia, es necesario entre los operadores B cercanos por la energía al operador A elegir el operador más fácilmente inversible.

¿Cómo construir operadores fácilmente inversibles? Es obvio, que si B^1, B^2, . . ., B^p son los operadores fácilmente inversibles, entonces el operador producto $B=B^1B^2\ldots B^p$, también es fácilmente inversible.

Señalemos que, a diferencia de los factores, el propio operador B puede tener una estructura compleja. Por ejemplo, sea $B^\alpha=E+\omega R_\alpha$, $\alpha=1, 2$, donde R_α es el operador correspondiente al operador de diferencias $(-\mathscr{R}_\alpha)$: $\mathscr{R}_\alpha y=y_{\bar{x}_\alpha x_\alpha}$, $\alpha=1, 2$. Al operador B^α le corresponde un operador de diferencias tripuntual, que se invierte por el método de factorización con un gasto del número de operaciones aritméticas, proporcional al número de incógnitas en el problema. El operador $B=B^1B^2$ tiene un modelo

nonapuntual y a él le corresponde el operador de diferencias $\mathscr{B}$:

$$\mathscr{B}y = y - \omega \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha x_\alpha} + \omega^2 y_{\bar{x}_1 x_1 \bar{x}_2 x_2}.$$

La complejidad de la estructura del operador B permite aumentar el cociente $\xi = \gamma_1/\gamma_2$, lo que conduce al aumento de la velocidad de convergencia del método iterativo.

Al construir el operador B se puede partir de un cierto operador $R = R^* > 0$ (regularizador), energéticamente equivalente a A y B:

$$c_1 R \leqslant A \leqslant c_2 R, \quad c_2 \geqslant c_1 > 0, \tag{5}$$

$$\overset{\circ}{\gamma}_1 B \leqslant R \leqslant \overset{\circ}{\gamma}_2 B, \quad \overset{\circ}{\gamma}_2 \geqslant \overset{\circ}{\gamma}_1 > 0. \tag{6}$$

Entonces son válidas las desigualdades (4) con las constantes $\gamma_1 = c_1 \overset{\circ}{\gamma}_1$ y $\gamma_2 = c_2 \overset{\circ}{\gamma}_2$, al mismo tiempo

$$\xi = \gamma_1/\gamma_2 = (c_1/c_2)\overset{\circ}{\xi}, \quad \overset{\circ}{\xi} = \overset{\circ}{\gamma}_1/\overset{\circ}{\gamma}_2.$$

¿En qué consiste el sentido de introducción del regularizador R? Para los problemas de contorno elípticos reticulares el operador R se elige, generalmente, de manera tal, que las constantes c_1 y c_2 en las desigualdades (5) no dependen de los parámetros de la red (del número de nodos de la red). Por ejemplo, si el operador A corresponde al operador de diferencias de coeficientes variables

$$\Lambda y = (a_1 y_{\bar{x}_1})_{x_1} + (a_2 y_{\bar{x}_2})_{x_2}, \quad 0 < c_1 \leqslant a_\alpha \leqslant c_2.$$

definido sobre la red uniforme $\bar{\omega} = \{x_{ij} = (ih_1, jh_2),\ 0 \leqslant i \leqslant N_1,\ 0 \leqslant j \leqslant N_2;\ h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2\}$, introducida en el rectangulo $\bar{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2\}$, de manera que $Ay = -\Lambda \overset{\circ}{y}$, donde $y(x) = \overset{\circ}{y}(x)$ para $x \in \omega$ y $\overset{\circ}{y}(x) = 0$ para $x \in \gamma$, entonces en calidad de R se puede tomar el operador correspondiente al operador de Laplace de diferencia $\mathscr{R}y = (\mathscr{R}_1 + \mathscr{R}_2)\, y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2}$, $Ry = -R\overset{\circ}{y}$, donde los operadores $\mathscr{R}_\alpha$ están definidos más arriba.

Empleando las fórmulas de Green de diferencias, es fácil mostrar (véase el punto 8 del § 2 del cap. V) que los operadores A y B son autoconjugados en H y se cumplen las desigualdades (5). Aquí H es el espacio de las funciones reticulares definidas sobre $\bar{\omega}$, en el cual el producto escalar se define por la fórmula $(u, v) = \sum_{x \in \bar{\omega}} u(x)\, v(x)\, h_1 h_2$.

Supongamos ahora que el operador A corresponde al operador elíptico de diferencias que contiene derivadas mixtas

$$\Lambda y = \sum_{\alpha,\,\beta=1}^{2} 0{,}5\,[(k_{\alpha\beta} y_{\bar{x}_\beta})_{x_\alpha} + (k_{\alpha\beta} y_{x_\beta})_{\bar{x}_\alpha}],$$

y que se cumplen las condiciones de elipticidad fuerte:

$$c_1 \sum_{\alpha=1}^{2} \xi_\alpha^2 \leqslant \sum_{\alpha,\,\beta=1}^{2} k_{\alpha\beta}(x)\,\xi_\alpha \xi_\beta \leqslant c_2 \sum_{\alpha=1}^{2} \xi_\alpha^2, \quad c_1 > 0.$$

Tomaremos en calidad de regularizador el operador R definido más arriba. En el punto 8 del § 2 del capítulo V fue mostrado, que para los operadores A y R examinados han sido cumplidas las desigualdades (5).

Citemos otro ejemplo. Supongamos que el operador A corresponde al operador de Laplace de diferencias de orden de exactitud elevado

$$\Lambda y = y_{\bar{x}_1 x_1} + y_{\bar{x}_2 x_2} + \frac{h_1^2 + h_2^2}{12}\, y_{\bar{x}_1 x_1 \bar{x}_2 x_2}.$$

Mostremos que si en calidad de operador R se elige el operador indicado más arriba, entonces son válidas las desigualdades (5) con las constantes $c_1 = 2/3$ y $c_2 = 1$.

Realmente, utilizando la primera fórmula de Green de diferencias y la igualdad $y_{\bar{x}_1 x_1 \bar{x}_2 x_2} = y_{\bar{x}_1 \bar{x}_2 x_1 x_2}$, válida para las funciones reticulares, definidas sobre la red rectangular $\overline{\omega}$, obtendremos

$$-(\Lambda \overset{\circ}{y},\ \overset{\circ}{y}) = (\overset{\circ}{y}_{\bar{x}_1}^2,\ 1)_1 + (\overset{\circ}{y}_{\bar{x}_2}^2,\ 1)_2 - \frac{h_1^2 + h_2^2}{12}\,(\overset{\circ}{y}_{\bar{x}_1 \bar{x}_2}^2,\ 1)_{12}, \tag{7}$$

$$-(\mathcal{R} \overset{\circ}{y},\ \overset{\circ}{y}) = (\overset{\circ}{y}_{\bar{x}_1}^2,\ 1)_1 + (\overset{\circ}{y}_{\bar{x}_2}^2,\ 1)_2.$$

Aquí está designado que

$$(u,\ v)_1 = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2 - 1} u(i,\ j)\, v(i,\ j)\, h_1 h_2,$$

$$(u,\ v)_2 = \sum_{i=1}^{N_1 - 1} \sum_{j=1}^{N_2} u(i,\ j)\, v(i,\ j)\, h_1 h_2,$$

$$(u,\ v)_{12} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} u(i,\ j)\, v(i,\ j)\, h_1 h_2.$$

De (7) se deduce la estimación $A \leqslant R$, es decir en (5) $c_2 = 1$. Luego, teniendo en cuenta que $\overset{\circ}{y}_{\overline{x}_2}(x) = 0$ para $x_1 = 0$, l_1 y $\overset{\circ}{y}_{\overline{x}_1} = 0$ para $x_2 = 0$, l_2, del lema 12 del capítulo V obtenemos la estimación

$$(\overset{\circ}{y}^2_{x_1\overline{x}_2},\ 1)_{12} \leqslant \frac{4}{h_2^2}(\overset{\circ}{y}^2_{\overline{x}_1},\ 1)_1 \tag{8}$$

y de manera análoga

$$(\overset{\circ}{y}^2_{x_1\overline{x}_2},\ 1)_{12} = (\overset{\circ}{y}^2_{x_2\overline{x}_1},\ 1)_{12} \leqslant \frac{4}{h_1^2}(\overset{\circ}{y}^2_{\overline{x}_2},\ 1)_2. \tag{9}$$

Multiplicando (8) por $h_2^2/12$, y (9), por $h_1^2/12$ y sumando las desigualdades obtenidas, tendremos

$$\frac{h_1^2+h_2^2}{12}(\overset{\circ}{y}^2_{x_1\overline{x}_2},\ 1)_{12} \leqslant \frac{1}{3}[(\overset{\circ}{y}^2_{\overline{x}_1},\ 1)_1 + (\overset{\circ}{y}^2_{\overline{x}_2},\ 1)_2].$$

De aquí y de (7) se sigue la estimación $A \geqslant 2R/3$. La afirmación está demostrada.

Los ejemplos examinados muestran, que para diferentes operadores A en calidad de regularizador puede ser elegido un mismo operador R. Por lo tanto el problema de construir un operador B para un esquema iterativo implícito se simplifica. El operador B se construye a partir de la condición de cercanía por la energía al regularizador R. La clase de los regularizadores es esencialmente más estrecha que la clase que contiene los operadores A. Si el operador B ya ha sido elegido y, por consiguiente, las constantes $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$ en las desigualdades (6) han sido halladas, entonces para cada operador concreto A quedan por hallar únicamente las constantes c_1 y c_2 en las desigualdades (5).

La dificultad principal al emplear un regularizador se traslada a la obtención de las estimaciones de $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$. Muy frecuentemente el operador B tiene una forma factorizada, y al mismo tiempo los factores dependen de ciertos parámetros iterativos. Por eso mismo prefija la familia de los operadores B de una estructura determinada, caracterizada por los parámetros indicados. Estos parámetros conviene elegir de la condición de máximo de $\overset{\circ}{\xi}$. Algunos ejemplos, en los cuales se estudian operadores factorizados, serán examinados en el punto siguiente. Aquí señalemos que en calidad de operador B a veces se puede tomar el regularizador R ($\overset{\circ}{\gamma}_1 = \overset{\circ}{\gamma}_2 = 1$).

2. **Esquemas iterativos del operador factorizado.** En el punto 1 fue ilustrado el principio de regularización mediante el ejemplo de un operador autoconjugado A. En este caso las desigualdades (4) son un corolario de las desigualdades (5) y (6).

En el capítulo VI fue mostrado, que si el operador A no es autoconjugado en H, y el espacio energético H_D está generado por el operador autoconjugado y definido positivo D, donde D es o bien B, ó $A^*B^{-1}A$, entonces las desigualdades (4) se deben cambiar por las desigualdades

$$\gamma_1 (Bx, x) \leqslant (Ax, x), \quad (B^{-1}Ax, Ax) \leqslant \gamma_2 (Ax, x),$$
$$\gamma_1 > 0 \tag{10}$$

ó por las desigualdades

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad (B^{-1}A_1 x, A_1 x) \leqslant \gamma_3^2 (Bx, x),$$
$$\gamma_1 > 0, \tag{11}$$

donde $A_1 = 0{,}5\,(A - A^*)$ es la parte no autoconjugada del operador A. Sea el operador $B = B^* > 0$ construido, partiendo del regularizador R, y supongamos que son cumplidas las desigualdades (6). Luego, si el operador R satisface las condiciones

$$c_1 (Rx, x) \leqslant (Ax, x), \quad (R^{-1}Ax, Ax) \leqslant c_2 (Ax, x),$$
$$c_1 > 0, \tag{10'}$$

entonces tienen lugar las desigualdades (10) con las constantes $\gamma_1 = c_1\gamma_1$, $\gamma_2 = c_2\gamma_2$.

En efecto, del lema 9 del capítulo V y de la desigualdad (6) se desprende, que son ciertas las desigualdades $\gamma_1 R^{-1} \leqslant \leqslant B^{-1} \leqslant \gamma_2 R^{-1}$. De aquí obtenemos

$$(B^{-1}Ax, Ax) \leqslant \gamma_2 (R^{-1}Ax, Ax) \leqslant c_2\gamma_2 (Ax, x).$$

Análogamente se demuestra, que si el operador R satisface las condiciones

$$c_1 R \leqslant A \leqslant c_2 R, \quad (R^{-1}A_1 x, A_1 x) \leqslant c_3^2 (Rx, x), \tag{11'}$$
$$c_1 > 0,$$

entonces son ciertas las desigualdades (11) con las constantes

$$\gamma_1 = c_1\gamma_1, \quad \gamma_2 = c_2\gamma_2 \quad \text{y} \quad \gamma_3 = c_3\gamma_3.$$

De esta forma, también en el caso de un operador no autoconjugado A es necesario poder obtener estimados para los $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$, que entran en las desigualdades (6).

Ocupémonos ahora de obtener las desigualdades (6) para los operadores autoconjugados R y B. Examinemos dos casos.

1) El operador R está representado en forma de la suma $R = R_1 + R_2$ de los operadores R_1 y R_2 conjugados uno del otro:

$$R_2 = R_1^*, \tag{12}$$

de manera que $(R_1x, x) = (R_2x, x) = 0{,}5\,(Rx, x)$, $x \in H$, y el operador B tiene la forma

$$B = (E + \omega R_1)(E + \omega R_2), \tag{13}$$

donde $\omega > 0$ es un parámetro.

2) El operador R lo representamos en forma de la suma $R = R_1 + R_2 + \ldots + R_p$, $p \geqslant 2$, de los operadores autoconjugados y conmutables de dos en dos R_α, $\alpha = 1, 2, \ldots, p$, así que

$$R_\alpha = R_\alpha^*, \quad R_\alpha R_\beta = R_\beta R_\alpha,$$
$$\alpha, \beta = 1, 2, \ldots, p, \tag{14}$$

y el operador B está factorizado y tiene la forma

$$B = \prod_{\alpha=1}^{p} (E + \omega R_\alpha), \tag{15}$$

donde $\omega > 0$ es un parámetro.

En cada uno de los casos el operador B es autoconjugado en H. Subrayemos especialmente la universalidad de la elección del operador B en la forma (13), donde los operadores R_1 y R_2 satisfacen la condición (12).

Nuestro problema consiste en la obtención de estimaciones para $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$, contenidas en (6), y en la elección del parámetro iterativo ω de la condición de máximo de la relación $\overset{\circ}{\xi} = \overset{\circ}{\gamma}_1/\overset{\circ}{\gamma}_2$.

Cada caso lo investigaremos por separado. El primer caso fue estudiado detalladamente en el capítulo X, dedicado al método alternativo triangular. Por eso nos limitaremos aquí solamente a la formulación de los resultados.

TEOREMA 1. *Sean cumplidas las condiciones (12) y en las desigualdades*

$$R \geqslant \delta E, \quad (R_2 x,\ R_2 x) \leqslant \frac{\Delta}{4}(Rx,\ x), \quad \delta > 0 \tag{16}$$

sean dadas las constantes δ *y* Δ. *Entonces para el valor óptimo del parámetro* $\omega = \omega_0 = 2/\sqrt{\delta\Delta}$ *el operador* B, *definido por la igualdad (13), satisface las desigualdades (6) con las constantes*

$$\overset{\circ}{\gamma}_1 = \frac{\delta}{2(1+\sqrt{\eta})}, \quad \overset{\circ}{\gamma}_2 = \frac{\delta}{4\sqrt{\eta}}, \quad \eta = \frac{\delta}{\Delta}.$$

Señalemos, que se puede examinar un tipo más general de operador B que (13), precisamente:

$$B = (\mathscr{D} + \omega R_1)\,\mathscr{D}^{-1}\,(\mathscr{D} + \omega R_2),$$

donde $\mathscr{D} = \mathscr{D}^* > 0$. Del lema 1 del capítulo X se deduce, que el teorema 1 se mantiene válido; es necesario solamente cambiar (16) por las desigualdades siguientes:

$$R \geqslant \delta\mathscr{D}, \quad (\mathscr{D}^{-1}R_2 x,\ R_2 x) \leqslant \frac{\Delta}{4}(Rx,\ x).$$

Aquí el operador $\mathscr{D}$ juega el papel de un parámetro iterativo complementario.

El operador B es fácilmente inversible, por ejemplo, en el caso, cuando al operador R_1 le corresponde una matriz triangular inferior, a R_2 una matriz triangular superior, y a $\mathscr{D}$ una matriz diagonal. Si el operador R corresponde a un operador elíptico de diferencias, entonces las matrices triangulares indicadas tendrán en cada fila un número finito de elementos no nulos, el cual no depende del número de nodos de la red. Por eso la inversión de cada factor, que entra en el operador B, pueda ser realizada con el gasto de un número de operaciones proporcional al número de incógnitas en el problema.

Examinemos ahora el segundo caso.

TEOREMA 2. *Supongamos que el operador* B *tiene la forma (15), se cumplen las condiciones (14) y están dadas las cotas de los operadores* R_α:

$$\delta_\alpha E \leqslant R_\alpha \leqslant \Delta_\alpha E, \quad \delta_\alpha > 0, \quad \alpha = 1,\ 2,\ \ldots,\ p.$$

Entonces para el valor óptimo del parámetro ω

$$\omega = \omega_0 = \frac{1}{\Delta}\,\frac{1-\eta^{1/p}}{\eta^{1/p}-\eta}$$

el operador B satisface las desigualdades (6) con las constantes

$$\overset{\circ}{\gamma}_1 = \frac{p\Delta}{(1+\omega_0\Delta)^p}, \quad \overset{\circ}{\gamma}_2 = \overset{\circ}{\gamma}_1 \frac{p-k(1-\eta)}{p\eta^{k/p}}, \quad \eta = \frac{\delta}{\Delta},$$

donde

$$\delta = \min_\alpha \delta_\alpha, \quad \Delta = \max_\alpha \Delta_\alpha, \quad k = \left[\frac{p}{1-\eta} - \frac{\eta^{1/p}}{1-\eta^{1/p}}\right],$$

y $[a]$ *es la parte entera del número* a.

En vista de lo voluminoso de la demostración nosotros no la citaremos. Señalemos únicamente que en virtud de las condiciones (14) el operador B es conmutable con los operadores R_α, $\alpha = 1, 2, \ldots, p$, y por eso

$$\overset{\circ}{\gamma}_1 = \min_{\delta \leqslant x_\alpha \leqslant \Delta} \frac{x_1 + x_2 + \ldots + x_p}{\prod_{\alpha=1}^{p} (1+\omega x_\alpha)},$$

$$\overset{\circ}{\gamma}_2 = \max_{\delta \leqslant x_\alpha \leqslant \Delta} \frac{x_1 + x_2 + \ldots + x_p}{\prod_{\alpha=1}^{p} (1+\omega x_\alpha)}.$$

Señalemos unos casos particulares del teorema 2. Si $p = 2$, entonces

$$k = 1, \quad \omega_0 = \frac{1}{\sqrt{\delta\Delta}}, \quad \overset{\circ}{\gamma}_1 = \frac{2\delta}{(1+\sqrt{\eta})^2},$$

$$\overset{\circ}{\gamma}_2 = \frac{\delta}{\sqrt{\eta}} \frac{1+\eta}{(1+\sqrt{\eta})^2}.$$

Si $p = 3$, entonces

$$k = 2, \quad \omega_0 = \frac{1}{\sqrt[3]{\delta\Delta}\,(\sqrt[3]{\delta} + \sqrt[3]{\Delta})},$$

$$\overset{\circ}{\gamma}_1 = 3\delta \left(\frac{1-\eta^{2/3}}{1-\eta}\right)^3, \quad \overset{\circ}{\gamma}_2 = \frac{\delta(1+2\eta)}{\eta^{2/3}} \left(\frac{1-\eta^{2/3}}{1-\eta}\right)^3.$$

Para el caso $p = 2$ se pueden obtener las mejores estimaciones para $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$, introduciendo en el operador

$$B = (E + \omega_1 R_1)(E + \omega_2 R_2) \tag{17}$$

dos parámetros ω_1 y ω_2, los cuales tienen en cuenta, que las cotas de los operadores R_1 y R_2 son diferentes. Tiene lugar

TEOREMA 5. *Supongamos que el operador B tiene la forma (17), se cumplen las condiciones*

$$R_\alpha = R_\alpha^*, \quad \alpha = 1, 2, \quad R_1 R_2 = R_2 R_1,$$

y están dadas las cotas de los operadores R_1 y R_2:

$$\delta_\alpha E \leqslant R_\alpha \leqslant \Delta_\alpha E, \quad \alpha = 1, 2, \quad \delta_1 + \delta_2 > 0.$$

Entonces para los valores óptimos de los parámetros ω_1 y ω_2

$$\omega_1 = \frac{1 + t\sqrt{\eta}}{r\sqrt{\eta} + s} \qquad \omega_1 = \frac{1 - t\sqrt{\eta}}{r\sqrt{\eta} - s},$$

se cumplen las desigualdades (6) con las constantes

$$\overset{\circ}{\gamma}_1 = \frac{4\sqrt{\eta}}{(\omega_1 + \omega_2)(1 + \sqrt{\eta})^2}, \qquad \overset{\circ}{\gamma}_2 = \frac{2(1+\eta)}{(\omega_1 + \omega_2)(1 + \sqrt{\eta})^2},$$

donde

$$r = \frac{\Delta_2 + \Delta_1 b}{1 + b}, \quad s = \frac{\Delta_2 - \Delta_1 b}{1 + b}, \quad t = \frac{1 - b}{1 + b}, \quad \eta = \frac{1 - a}{1 + a},$$

$$a = \sqrt{\frac{(\Delta_1 - \delta_1)(\Delta_2 - \delta_2)}{(\Delta_1 + \delta_2)(\Delta_2 + \delta_1)}}, \quad b = \frac{\Delta_2 + \delta_1}{\Delta_1 - \delta_1} a.$$

Para la demostración del teorema realicemos un cambio, poniendo

$$R_1 = (r\overline{R}_1 - sE)(E - t\overline{R}_1)^{-1}, \quad R_2 = (r\overline{R}_2 + sE)(E + t\overline{R}_2)^{-1}$$

con los valores indicados para r, s y t. Se puede mostrar, que los operadores definidos de esta forma

$$\overline{R}_1 = (R_1 + sE)(rE + tR_1)^{-1}, \qquad \overline{R}_2 = (R_2 - sE)(rE - tR_2)^{-1}$$

satisfacen las condiciones $\overline{R}_\alpha = \overline{R}_\alpha^*$, $\alpha = 1, 2$, $\overline{R}_1 \overline{R}_2 = \overline{R}_2 \overline{R}_1$ y tienen las mismas cotas $\eta E \leqslant \overline{R}_\alpha \leqslant E$, $\eta > 0$, $\alpha = 1, 2$. Después, como los operadores $E - t\overline{R}_1$ y $E + t\overline{R}_2$ son autoconjugados y definidos positivos, entonces existen los operadores conmutativos $(E - t\overline{R}_1)^{1/2}$ y $(E + t\overline{R}_2)^{1/2}$. Pongamos

$$x = (E - t\overline{R}_1)^{1/2}(E + t\overline{R}_2)^{1/2} y.$$

Obtendremos

$$(Bx, x) = (1 - \omega_1 s)(1 + \omega_2 s)(\overline{B}y, y), \tag{18}$$

$$(Rx, x) = (r - st)(\overline{R}y, y), \tag{19}$$

donde $\overline{B} = (E + \overline{\omega}\overline{R}_1)(E + \overline{\omega}\overline{R}_2)$, $\overline{R} = \overline{R}_1 + \overline{R}_2$,

$$\overline{\omega} = \frac{\omega_1 r - t}{1 - \omega_1 s} = \frac{\omega_2 r + t}{1 + \omega_2 s}. \tag{20}$$

De (20) hallaremos

$$2\bar{\omega} = \frac{\omega_1 r - t}{1 - \omega_1 s} + \frac{\omega_2 r + t}{1 + \omega_2 s} = \frac{(r - st)(\omega_1 + \omega_2)}{(1 - \omega_1 s)(1 + \omega_2 s)}.$$

De aquí y de (18), (19) obtenemos

$$\frac{(Rx,\ x)}{(Bx,\ x)} = \frac{2\bar{\omega}}{\omega_1 + \omega_2} \frac{(\bar{R}y,\ y)}{(\bar{B}y,\ y)}. \tag{21}$$

Utilizando el teorema 2, obtendremos, que para

$$\bar{\omega} = \omega_0 = 1/\sqrt{\eta} \tag{22}$$

tienen lugar las desigualdades

$$\overset{\circ}{\gamma}_1 (\bar{R}y,\ y) \leqslant \bar{B}y,\ y) \leqslant \overset{\circ}{\gamma}_2 (\bar{R}y,\ y), \tag{23}$$

donde

$$\overset{\circ}{\gamma}_1 = \frac{2\eta}{(1 + \sqrt{\eta})^2},\quad \overset{\circ}{\gamma}_2 = \frac{\sqrt{\eta}(1 + \eta)}{(1 + \sqrt{\eta})^2}.$$

Por consiguiente, de (20) y (22) obtendremos los valores óptimos para los parámetros ω_1 y ω_2:

$$\omega_1 = \frac{1 + t\sqrt{\eta}}{r\sqrt{\eta} + s},\quad \omega_2 = \frac{1 - t\sqrt{\eta}}{r\sqrt{\eta} - s},$$

y de (21) y (23) se deducen las desigualdades (6) con las constantes $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$, indicadas en la formulación del teorema 3. El teorema 3 está demostrado.

3. Procedimiento de inversión implícita del operador B (método bictápico). En el punto 2 nosotros estudiamos un procedimiento para construir métodos iterativos implícitos que se caracteriza por lo que el operador B se prefija constructivamente en la forma de un producto de operadores fácilmente inversibles. Examinemos otro método, en el cual la aproximación iterativa y_{k+1} se encuentra como resultado de un procedimiento auxiliar el cual se puede interpretar como la inversión implícita de algún operador B.

Recordemos, que la idea general de este procedimiento está examinada en el punto 4 del § 3 del cap. V. En el punto 4 del § 1 del cap. XIII este método se utilizó durante la construcción del método iterativo de resolución de una ecuación con un operador no lineal A. Allí mismo fueron formuladas las condiciones que permiten obtener las estimaciones para $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$ entrantes en las desigualdades (6).

Expongamos los resultados obtenidos. Supongamos que la aproximación iterativa y_{k+1} se encuentra por la fórmula del esquema con corrección: $y_{k+1} = y_k - \tau_{k+1} w^p$, y la corrección w^p es una aproximación de la ecuación auxiliar

$$Rw = r_k, \quad r_k = Ay_k - f. \tag{24}$$

Aquí R es un regularizador que satisface las desigualdades (5) para el caso de un operador autoconjugado A y satisface las desigualdades (10′) ó (11′) para A no autoconjugado.

Supongamos que la ecuación (24) se resuelve con ayuda de algún esquema iterativo de dos capas, de manera que el error $z^m = w^m - w$ satisfaga la ecuación

$$z^{m+1} = S_{m+1} z^m, \quad m = 0, 1, \ldots, p-1, \quad z^0 = w^0 - w,$$

donde S_{m+1} es el operador de transición de la m-ésima a la $(m+1)$-ésima iteración.

Eligiendo $w^0 = 0$, de las igualdades

$$z^p = w^p - w = T_p (w^0 - w), \quad T_p = \prod_{m=1}^{p} S_m,$$

$$w = R^{-1} r_k,$$

tendremos

$$w^p = B^{-1} r_k, \quad \text{donde} \quad B = R (E - T_p)^{-1}.$$

Sustituyendo la expresión hallada para w^p en (23), obtenemos el esquema iterativo implícito (2) con el operador B indicado.

TEOREMA 4. *Sean cumplidas las condiciones*

$$R = R^* > 0, \quad T_p^* R = R T_p, \quad \| T_p \|_R \leqslant q < 1.$$

Entonces el operador $B = R (E - T_p)^{-1}$ *es autoconjugado y definido positivo en* H *y son ciertas las desigualdades* (6) *con las constantes* $\mathring{\gamma}_1 = 1 - q$ *y* $\mathring{\gamma}_2 = 1 + q$.

Para la demostración véase el lema 2 del capítulo XIII.

OBSERVACIÓN. Si los operadores R y T_p son autoconjugados y conmutables y $\| T_p \| \leqslant q < 1$, entonces son válidas las afirmaciones del teorema 4.

Por el procedimiento descrito más arriba nosotros construímos un esquema iterativo implícito de dos capas. Si partimos de las fórmulas

$$y_{k+1} = \alpha_{k+1} y_k + (1 - \alpha_{k+1}) y_{k-1} - \tau_{k+1} \alpha_{k+1} w_k^p,$$

$$k = 1, 2, \ldots,$$

$$y_1 = y_0 - \tau_1 w_0^p,$$

y encontramos la corrección w_k^p para cualquier $k = 0, 1, \ldots$ como solución aproximada de la ecuación (24), entonces nosotros obtendremos el esquema iterativo implícito de tres capas

$$By_{k+1} = \alpha_{k+1}(B - \tau_{k+1}A)\, y_k + (1 - \alpha_{k+1})\, By_{k-1} + \\ + \alpha_{k+1}\tau_{k+1}f, \qquad (25)$$

$$By_1 = (B - \tau_1 A)\, y_0 + \tau_1 f.$$

Como conclusión señalemos que los parámetros iterativos τ_k para el esquema (2) y τ_k, α_k para el esquema (25) se eligen de acuerdo a la teoría general de los métodos iterativos. Aquí aparece el problema de la elección de un número óptimo de iteraciones p para el proceso iterativo auxiliar. Expliquemos la situación. Supongamos para simplificar que el proceso auxiliar es estacionario ($S_m \equiv S$), los operadores R y S autoconjugados y conmutables y se cumple la condición $\| S \| \leqslant \rho$. Entonces $q = \rho^p$, es decir,

$$p = \ln q / \ln \rho. \qquad (26)$$

Los operadores A y B satisfacen las desigualdades (4) con las constantes

$$\gamma_1 = c_1 (1 - q), \quad \gamma_2 = c_2 (1 + q).$$

Si los parámetros iterativos τ_k en el esquema (2) son elegidos por las fórmulas del método de Chebishev, entonces para el número de iteraciones es cierta la estimación

$$n \geqslant n_0(\varepsilon), \quad n_0(\varepsilon) = \ln(0{,}5\varepsilon)/\ln \rho_1,$$

donde

$$\rho_1 = \rho_1(q) = \frac{1 - \sqrt{\xi}}{1 + \sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2} = \frac{c_1}{c_2}\,\frac{1-q}{1+q}.$$

Entonces el número total de iteraciones $k = pn$ se estima por la cantidad

$$k \geqslant k_0(\varepsilon), \quad k_0(\varepsilon) = \frac{\ln 0{,}5\varepsilon}{\ln \rho}\,\frac{\ln p}{\ln \rho_1(q)}.$$

De aquí se deduce que la magnitud q, la cual determina de acuerdo con (26), el número de iteraciones interiores, debe ser elegida de la condición de mínimo de la función $\varphi(q) = \ln q / \ln \rho_1(q)$. Este problema puede ser resuelto numéricamente.

§ 2. Ejemplos des resolución de los problemas de contorno elípticos

1. Métodos directos e iterativos. Examinemos los métodos de resolución de los problemas multidimensionales de diferencias de contorno para las ecuaciones elípticas. Comenzamos por el problema de Dirichlet para la ecuación de Poisson en un paralelepípedo rectangular p-dimensional $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2, \ldots, p,\ p \geqslant 3\}$

$$Lu = \sum_{\alpha=1}^{p} \frac{\partial^2 u}{\partial x_\alpha^2} = -f(x), \quad x = (x_1, \ldots, x_p) \in G, \qquad (1)$$
$$u(x) = g(x), \quad x \in \Gamma.$$

En la red rectangular uniforme $\overline{\omega} = \{x_i = (i_1 h_1, \ldots, i_p h_p) \in \overline{G},\ 0 \leqslant i_\alpha \leqslant N_\alpha,\ h_\alpha N_\alpha = l_\alpha,\ \widetilde{\alpha} = 1, 2, \ldots, p\}$ al problema (1) le ponemos en correspondencia el problema de Dirichlet de diferencias

$$\Lambda y = \sum_{\alpha=1}^{p} \Lambda_\alpha y = -\varphi(x), \quad x \in \omega$$
$$y(x) = g(x), \qquad x \in \gamma \qquad (2)$$

donde $\Lambda_\alpha y = y_{\overline{x}_\alpha x_\alpha}$, $1 \leqslant \alpha \leqslant p$.

Para resolver los problemas (2) hacemos uso del método de separación de variables. Lo mismo que en el caso de $p = 2$ se puede mostrar que el problema de valores propios para el operador p-dimensional de Laplace de diferencias sobre la red $\overline{\omega}$

$$\Lambda\mu + \lambda\mu = 0, \quad x \in \omega, \quad \mu(x) = 0, \quad x \in \gamma$$

tiene la siguiente solución:

$$\mu_k(i_1, i_2, \ldots, i_p) = \prod_{\alpha=1}^{p} \mu_{k_\alpha}^{\alpha}(i_\alpha),$$

$$\mu_{k_\alpha}^{\alpha}(i_\alpha) = \sqrt{\frac{2}{l_\alpha}} \operatorname{sen} \frac{k_\alpha \pi i_\alpha}{N_\alpha},$$

$$1 \leqslant k_\alpha \leqslant N_\alpha - 1, \quad 0 \leqslant i \leqslant N_\alpha, \quad 1 \leqslant \alpha \leqslant p,$$

$$\lambda_k = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{k_\alpha \pi}{2N_\alpha}, \quad k = (k_1, k_2, \ldots, k_p).$$

Las funciones $\mu_k(i_1, i_2, \ldots, i_p)$ son ortonormalizadas en el sentido del producto escalar

$$(u, v) = \sum_{i_1=1}^{N_1-1} \sum_{i_2=1}^{N_2-1} \cdots \sum_{i_p=1}^{N_p-1} u(i_1, i_2, \ldots, i_p) \times$$

$$\times v(i_1, i_2), \ldots, i_p) \prod_{\alpha=1}^{p} h_\alpha.$$

Además, toda función reticular $v(i_1, i_2, \ldots, i_p)$ que se anula en γ, puede ser representada en forma de la suma

$$v(i_1, i_2, \ldots i_p) = \sum_{k_1=1}^{N_1-1} \sum_{k_2=1}^{N_2-1} \cdots$$

$$\cdots \sum_{k_p=1}^{N_p-1} c_{k_1 k_2 \ldots k_p} \mu_k(i_1, i_2, \ldots, i_p),$$

donde los coeficientes de Fourier $c_{k_1 k_2 \ldots k_p}$ se determinan por la fórmula

$$c_{k_1 k_2 \ldots k_p} = \sum_{i_1=1}^{N_1-1} \sum_{i_2=1}^{N_2-1} \cdots \sum_{i_p=1}^{N_p-1} v(i_1, i_2, \ldots, i_p) \times$$

$$\times \mu_k(i_1, i_2, \ldots, i_p) \prod_{\alpha=1}^{p} h_\alpha.$$

A fin de emplear el método de separación de variables, hay que reducir el problema (2) a un problema con condiciones homogéneas de frontera. Para eso es suficiente cambiar en los nodos de contorno el segundo miembro $\varphi(x)$ de la ecuación (2). Vamos a considerar que esto ha sido hecho y en la ecuación (2) la función $g(x)$ es nula.

La solución del problema

$$\sum_{\alpha=1}^{p} y_{\bar{x}_\alpha x_\alpha} = -\varphi(x), \quad x \in \omega, \quad y(x) = 0, \quad x \in \gamma \tag{3}$$

la buscaremos en forma de la suma

$$y(i_1, i_2, \ldots, i_p) = \sum_{k_1=1}^{N_1-1} \sum_{k_2=2}^{N_2-1} \cdots$$

$$\cdots \sum_{k_p=1}^{N_p-1} c_{k_1 k_2 \ldots k_p} \mu_k(i_1, i_2, \ldots, i_p). \tag{4}$$

Desarrollando el segundo miembro $\varphi(x)$ en una suma análoga, de (3) obtendremos que tienen lugar las igualdades

$$c_{k_1 k_2 \ldots k_p} = \frac{\varphi_{k_1 k_2 \ldots k_p}}{\lambda k}, \quad 1 \leqslant k_\alpha \leqslant N_\alpha - 1, \quad 1 \leqslant \alpha \leqslant P, \qquad (5)$$

donde $\varphi_{k_1 k_2 \ldots k_p}$ son los coeficientes de Fourier de la función $\varphi(x)$, los cuales se calculan mediante la fórmula

$$\varphi_{k_1 k_2 \ldots k_p} = \sum_{i_1=1}^{N_1-1} \sum_{i_2=1}^{N_2-1} \cdots \sum_{i_p=1}^{N_p-1} \varphi(i_1, i_2, \ldots, i_p) \times$$

$$\times \mu_k(i_1, i_2, \ldots, i_p) \prod_{\alpha=1}^{p} h_\alpha. \qquad (6)$$

Así pues, primero para $1 \leqslant k_\alpha \leqslant N_\alpha - 1$, $1 \leqslant \alpha \leqslant p$ se calculan los coeficientes de Fourier $\varphi_{k_1 k_2 \ldots k_p}$, luego por la fórmula (5) se hallan $c_{k_1 k_2 \ldots k_p}$ y por último, para $1 \leqslant i_\alpha \leqslant N_\alpha - 1$, $1 \leqslant \alpha \leqslant p$, se determina la solución buscada $y(i_1, i_2, \ldots, i_p)$ mediante la fórmula (4).

Tomando en consideración la estructura de las funciones $\mu_k(i_1, i_2, \ldots, i_p)$, se puede proponer el algoritmo siguiente:

1) para $1 \leqslant i_\alpha = N_\alpha - 1$, $2 \leqslant \alpha \leqslant p$ se calculan las sumas

$$\varphi_{k_1}(i_2, i_3, \ldots, i_p) =$$

$$= \sum_{i_1=1}^{N_1-1} \varphi(i_1, i_2, \ldots, i_p) \operatorname{sen} \frac{\pi k_1 i_1}{N_1}, \quad 1 \leqslant k_1 \leqslant N_1 - 1,$$

2) sucesivamente, adoptando $m = 2, 3, \ldots, p-1$, se calculan para

$$1 \leqslant i_\alpha \leqslant N_\alpha - 1, \quad m+1 \leqslant \alpha \leqslant p$$

$$\text{y} \quad 1 \leqslant k_\beta \leqslant N_\beta - 1, \quad 1 \leqslant \beta \leqslant m-1$$

las sumas

$$\varphi_{k_1 k_2 \ldots k_m}(i_{m+1}, \ldots, i_p) =$$

$$= \sum_{i_m=1}^{N_m-1} \varphi_{k_1 k_2 \ldots k_{m-1}}(i_m, \ldots, i_p) \operatorname{sen} \frac{\pi k_m i_m}{N_m}$$

siendo $1 \leqslant k_m \leqslant N_m - 1$,

3) para $1 \leqslant k_\beta \leqslant N_\beta - 1$, $1 \leqslant \beta \leqslant p-1$ se determinan las sumas

$$\overline{c}_{k_1 k_2 \ldots k_p} = \frac{1}{\lambda_k} \prod_{\alpha=1}^{p} \frac{2}{N_\alpha} \sum_{i_p=1}^{N_p-1} \varphi_{k_1 k_2 \ldots k_{p-1}}(i_p) \operatorname{sen} \frac{\pi k_p i_p}{N_p}$$

siendo $1 \leqslant k_p \leqslant N_{p-1}$, y como resultado serán hallados (con exactitud hasta un factor) los coeficientes $c_{k_1 k_2 \ldots k_p}$,

4) para $1 \leqslant k_\beta \leqslant N_\beta - 1$, $2 \leqslant \beta \leqslant p$ se calculan las sumas

$$y_{k_\beta k_{\beta+1} \ldots k_p}(i_1) = \sum_{k_1=1}^{N_1-1} \bar{c}_{k_1 k_\beta \ldots k_p} \operatorname{sen} \frac{\pi k_1 i_1}{N_1}, \quad 1 \leqslant i_1 \leqslant N_1 - 1,$$

5) de una manera consecutiva, fijando $m = 2, 3, \ldots, p-1$, se calculan para $1 \leqslant k_\beta \leqslant N_\beta - 1$, $m+1 \leqslant \beta \leqslant p$ y $1 \leqslant i_\alpha \leqslant N_\alpha - 1$, $1 \leqslant \alpha \leqslant m-1$ las sumas

$$y_{k_{m+1} \ldots k_p}(i_1, i_2, \ldots i_m) = \sum_{k_m=1}^{N_m-1} y_{k_m \ldots k_p}(i_1, \ldots, i_{m-1}) \operatorname{sen} \frac{\pi k_m i_m}{N_m}$$

cuando $1 \leqslant i_m \leqslant N_m - 1$,

6) para $1 \leqslant i_\alpha \leqslant N_\alpha - 1$, $1 \leqslant \alpha \leqslant p-1$ se calculan las sumas

$$y(i_1, i_2, \ldots, i_p) = \sum_{k_p=1}^{N_p-1} y_{k_p}(i_1, i_2, \ldots, i_{p-1}) \operatorname{sen} \frac{\pi k_p i_p}{N_p}, \quad 1 \leqslant i_p \leqslant N_p - 1.$$

Ahora estimemos el número de operaciones aritméticas que son necesarias para realizar el algoritmo propuesto. Si $N_\alpha = 2^{n_\alpha}$, $n_\alpha \geqslant 1$, entonces para calcular las sumas simples indicadas se puede utilizar el algoritmo de la transformación discreta rápida de Fourier, en el cual con el fin de calcular las sumas del tipo siguiente

$$y_i = \sum_{k=1}^{N-1} a_k \operatorname{sen} \frac{\pi k i}{N}, \quad 1 \leqslant i \leqslant N-1, \quad N = 2^n$$

se necesitan (véase el punto 2 del § 1 del cap. IV) $q = (2N-1)\log_2 N - 3N + 3$ operaciones aritméticas. De aquí se deduce que el número total de operaciones requerido para calcular todas las sumas indicadas en los puntos 1) . . . 6)

se estima por la magnitud

$$Q=2\sum_{m=1}^{p}\left(2\log_2 N_m-3-\frac{\log N_m-3}{N_m}\right)\times$$

$$\times N_1N_2\ldots N_p\leqslant 4\prod_{\alpha=1}^{p}N_\alpha\log_2\left(\prod_{\alpha=1}^{p}N_\alpha\right).$$

Si $N_1=N_2=\ldots=N_p=N$, entonces $Q\approx 4pN^p\log_2 N$. A esta cantidad conviene añadir además las N^p operaciones de multiplicación que se efectúan en el punto 3).

Notemos que para $p=2$ el algoritmo construido aquí coincide con el algoritmo descrito en el p. 2 del § 2 del cap. IV.

Supongamos que ahora es menester resolver una ecuación elíptica de coeficientes variables para el problema de Dirichlet

$$Lu=\sum_{\alpha=1}^{p}\frac{\partial}{\partial x_\alpha}\left(k_\alpha(x)\frac{\partial u}{\partial x_\alpha}\right)=-f(x),\quad x\in G,$$
$$u(x)=\quad g(x),\quad x\in\Gamma. \tag{7}$$

En la red rectangular uniforme $\overline{\omega}$, introducida arriba, construyamos el problema de Dirichlet de diferencias (2), donde

$$\Lambda_\alpha y=(a_\alpha(x)\,y_{\overline{x}_\alpha})_{x_\alpha},\quad a_\alpha(x)=$$
$$=k_\alpha(x)\,|_{x_\alpha=(i_\alpha-0,5)h_\alpha},\qquad 1\leqslant\alpha\leqslant p.$$

Para la solución aproximada del problema (2) utilizaremos el método iterativo. Como siempre, reduzcamos el problema (2) a la ecuación operacional

$$Au=f \tag{8}$$

en el espacio de Hilbert H de las funciones reticulares, prefijadas en ω, en el cual el producto escalar se define con ayuda de la fórmula

$$(u,\ v)=\sum_{x\in\omega}u(x)\,v(x)\prod_{\alpha=1}^{p}h_\alpha.$$

Si se supone que los coeficientes $k_\alpha(x)$ satisfacen las condiciones

$$0<c_1\leqslant k_\alpha(x)\leqslant c_2,\quad x\in\overline{G},\quad \alpha=1,\ 2,\ \ldots,\ p, \tag{9}$$

entonces, aplicando las fórmulas de Green de diferencias, se puede mostrar que tienen lugar las desigualdades

$$c_1 R \leqslant A \leqslant c_2 R, \quad c_1 > 0 \tag{10}$$

donde el operador R corresponde al operador de Laplace p-dimensional de diferencias $(-\mathscr{R})$

$$\mathscr{R} y = \sum_{\alpha=1}^{p} y_{\bar{x}_\alpha x_\alpha}.$$

Además, los operadores A y R son autoconjugados en H.

A fin de resolver las ecuaciones (8) emplearemos un esquema iterativo implícito

$$B \frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = f, \quad k = 0, 1, \ldots, \quad y_0 \in H \tag{11}$$

con el conjunto de Chebishev de parámetros τ_k y el operador $B = R$. Para el operador indicado B el método iterativo (11) converge en H_D, donde D es ó bien A, ó B, ó $AB^{-1}A$, y para el número de iteraciones es válida la estimación $n \geqslant n_0(\varepsilon)$, donde

$$n_0(\varepsilon) = \frac{\ln 0{,}5\varepsilon}{\ln \rho_1}, \quad \rho_1 = \frac{1 - \sqrt{\xi}}{1 + \sqrt{\xi}}, \quad \xi = \frac{c_1}{c_2}.$$

A fin de encontrar y_{k+1} es necesario resolver el problema de Dirichlet para la ecuación de Poisson. Si con este objeto se emplea el método directo de separación de variables descrito más arriba, entonces el número total de operaciones aritméticas necesarias para resolver aproximadamente el problema (2) se estima por la magnitud $Q = 2pN^p \sqrt{\frac{c_2}{c_1}} \ln(2/\varepsilon) \log_2 N$ para la red cuadrada $(N_1 = N_2 = \ldots = N_p = N)$.

De modo análogo hay que proceder, cuando se necesita resolver el problema de Dirichlet para la ecuación elíptica que contiene derivadas mixtas

$$Lu = \sum_{\alpha,\, \beta=1}^{p} \frac{\partial}{\partial x_\alpha} \left(k_{\alpha\beta}(x) \frac{\partial u}{\partial x_\beta} \right) = -f(x),$$

cuyos coeficientes satisfacen la condición de elipticidad fuerte

$$c_1 \sum_{\alpha=1}^{p} \xi_\alpha^2 \leqslant \sum_{\alpha,\, \beta=1}^{p} k_{\alpha\beta}(x) \xi_\alpha \xi_\beta \leqslant c_2 \sum_{\alpha=1}^{p} \xi_\alpha^2, \quad c_1 > 0.$$

Examinemos ahora los ejemplos de los problemas de contorno elípticos, para la resolución aproximada de los cuales es necesario utilizar el esquema iterativo (11) con el operador factorizado B.

Supongamos que el problema (7) se debe resolver sobre una red arbitraria no uniforme $\bar{\omega}$, introducida en el paralelepípedo $\bar{G}$. El respectivo esquema de diferencias (2) con los operadores

$$\Lambda_\alpha y = (\alpha_\alpha(x)\, y_{\bar{x}_\alpha})_{\hat{x}_\alpha},\ 1 \leqslant \alpha \leqslant p,$$

se escribe en forma de la ecuación operacional (8), mientras que el producto escalar sobre H se prefija mediante la fórmula

$$(u, v) = \sum_{i_1=1}^{N_1-1} \sum_{i_2=1}^{N_2-1} \ldots \sum_{i_p=1}^{N_p-1} u(i)\, v(i)\, \hbar_1(i_1)\, \hbar_2(i_2) \ldots \hbar_p(i_p).$$

Como regularizador eligimos el operador $R = R_1 + R_2 + \ldots + R_p$, el cual corresponde al operador de Laplace de diferencias $(-\mathscr{R})$ sobre la red no uniforme $\bar{\omega}$

$$\mathscr{R} = \mathscr{R}_1 + \mathscr{R}_2 + \ldots + \mathscr{R}_p,\ \mathscr{R}_\alpha y = y_{\bar{x}_\alpha \hat{x}_\alpha},$$

$$\alpha = 1,\ 2,\ \ldots,\ p.$$

Aplicando las fórmulas de Green de diferencias cercioremonos de que los operadores R_α son autoconjugados en H. Además, los últimos son conmutables de dos en dos. De la observación 1 al lema 16 del cap. V se sigue que tiene lugar una desigualdad operacional

$$R_\alpha \geqslant \delta_\alpha E,\ \delta_\alpha = \frac{8}{l_\alpha^2},\quad \alpha = 1, 2, \ldots, p,$$

que da la cota inferior del operador R_α. La estimación de la cota superior del operador R_α

$$R_\alpha \leqslant \Delta_\alpha E,\ \Delta_\alpha = \max_{1 \leqslant i_\alpha \leqslant N_\alpha - 1} \frac{4}{h_\alpha(i_\alpha)\, h_\alpha(i_\alpha + 1)},$$

$$\alpha = 1,\ 1,\ \ldots,\ p$$

se desprende del lema 17 del cap. V.

Si se cumplen las condiciones (9), serán ciertas las desigualdades (10). Para resolver la ecuación (8) utilizaremos

el esquema iterativo (11), en que B es un operador factorizado

$$B = \prod_{\alpha=1}^{p} (E + \omega R_\alpha).$$

Del teorema 2 de este capítulo obtenemos el valor óptimo del parámetro iterativo ω, así como las constante $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$ en las desigualdades $\overset{\circ}{\gamma}_1 B \leqslant R \leqslant \overset{\circ}{\gamma}_2 B$. De aquí y de (10) se hallan $\gamma_1 = c_1\overset{\circ}{\gamma}_1$ y $\gamma_2 = c_2\overset{\circ}{\gamma}_2$ que son las constantes de la equivalencia energética de los operadores A y B. Empleándolas se puede construir para el esquema (11) el conjunto de Chébishev de los parámetros iterativos τ_k.

2. **Problema de Dirichlet de diferencias con elevado orden de exactitud en el caso multidimensional.** En este punto examinaremos los métodos de resolución de dicho problema para la ecuación de Poisson dada en el paralelepípedo rectangular p-dimensional. Recordemos que para $p = 2$ en el p. 2 del § 3 del cap. III y en el p. 3 del § 3 del cap. IV fueron construidos los métodos directos, y en el p. 3 del § 2 del cap. XI, el método iterativo de las direcciones variables.

Ahora supongamos que $p \geqslant 3$. Sobre el paralelepípedo p-dimensional $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha,\ \alpha = 1, 2, \ldots, p\}$ en la red $\overline{\omega} = \{x_i = (i_1h_1, \ldots, i_p h_p) \in \overline{G},\ 0 \leqslant i_\alpha \leqslant N_\alpha,\ h_\alpha N_\alpha = l_\alpha,\ \alpha = 1, 2, \ldots, p\}$ al problema (1) le ponemos en correspondencia un esquema de diferencias del cuarto orden de aproximación:

$$\Lambda y = \sum_{\alpha=1}^{p} \Lambda_\alpha + \sum_{\alpha=1}^{p} \varkappa_\alpha \sum_{\beta\neq\alpha}^{1\div p} (\Lambda_\alpha\Lambda_\beta)\, y = -\varphi(x), \quad x \in \omega,$$
$$y(x) = g(x), \quad x \in \gamma, \tag{12}$$

donde

$$\Lambda_\alpha y = y_{\bar{x}_\alpha x_\alpha},\ 1 \leqslant \alpha \leqslant p,\ \varphi = f + \sum_{\alpha=1}^{p} \varkappa_\alpha \Lambda_\alpha f,\ \varkappa_\alpha = \frac{h_\alpha^2}{12}.$$

Buscaremos la solución del problema (12) recurriendo al método de separación de variables. Como de costumbre, vamos a considerar que $g(x) \equiv 0$, de otro modo cambiaremos el segundo miembro $\varphi(x)$ de la ecuación (12) en los nodos de frontera, debido a lo que tendremos en cuenta la heterogeneidad de la condición de contorno.

Indiquemos que las funciones $\mu_k(i_1, i_2, \ldots, i_p)$, citadas en el p. 1 son funciones propias del operador de diferencias

Λ definido en (12), mientras que los números

$$\lambda_h = \sum_{\alpha=1}^{p} \lambda_{h_\alpha}^{(\alpha)} - \sum_{\alpha=1}^{p} \varkappa_\alpha \sum_{\beta \neq \alpha}^{1 \div p} \lambda_{h_\alpha}^{(\alpha)} \lambda_{h_\beta}^{(\beta)}, \quad h = (h, h_2, \ldots, h_p), \tag{13}$$

donde $\lambda_{h_\alpha}^{(\alpha)} = \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{k_\alpha \pi h_2}{2l_\alpha}$, $1 \leqslant k_\alpha \leqslant N_2 - 1$, $1 \leqslant \alpha \leqslant p$,

son los valores propios de este operador.

Desarrollando la solución $y(x)$ y el segundo miembro $\varphi(x)$ en una suma respecto a las funciones $\mu_h(i_1, i_2, \ldots, i_p)$ (como por lo común se hace en el método de separación de variables), obtendremos

$$y(i_1, i_2, \ldots, i_p) = \sum_{h_1=1}^{N_1-1} \sum_{h_2=1}^{N_2-1} \ldots \sum_{h_p=1}^{N_p-1} \frac{\varphi_{h_1 h_2 \ldots h_p}}{\lambda_h} \mu_h(i_1, i_2, \ldots, i_p),$$

donde $\varphi_{h_1 h_2 \ldots h_p}$ son los coeficientes de Fourier de la función $\varphi(x)$, definidos en (6).

El algoritmo de cálculo de semejantes sumas fue descrito con detalles en el punto 1. Comparándolo con el método para resolver el problema de Dirichlet de diferencias con exactitud de segundo orden, construido en el punto 1, aquí aparece sólo una diferencia, a saber, otros valores propios de λ_h. De ese modo, dichos métodos prácticamente son equivalentes en cuanto a los gastos para el trabajo de cálculo. Al mismo tiempo, el esquema de diferencias (12) asegurará una exactitud de resolución del problema diferencial (1) más alta que el esquema (2).

Notemos que en virtud de las desigualdades

$$\lambda_{h_\alpha}^{(\alpha)} < \frac{4}{h_\alpha^2}, \quad 1 \leqslant k_\alpha \leqslant N_\alpha - 1, \quad 1 \leqslant \alpha \leqslant p,$$

de (13) tendremos

$$\lambda_h \geqslant \sum_{\alpha=1}^{p} \lambda_{h_\alpha}^{(\alpha)} - \frac{1}{3} \sum_{\alpha=1}^{p} \sum_{\beta \neq \alpha}^{1 \div p} \lambda_{h_\beta}^{(\beta)} =$$

$$= \sum_{\alpha=1}^{p} \lambda_{h_\alpha}^{(\alpha)} - \frac{1}{3} \sum_{\alpha=1}^{p} \left(\sum_{\beta=1}^{p} \lambda_{h_\beta}^{(\beta)} - \lambda_{h_\alpha}^{(\alpha)} \right) = \frac{4-p}{3} \sum_{\alpha=1}^{p} \lambda_{h_\alpha}^{(\alpha)}.$$

Cuando $p \geqslant 4$, el operador de diferencias Λ que determina el esquema de diferencias (12) pierde la propiedad de signo definido (elipticidad).

Sin alterar el orden de aproximación, se puede elegir otro operador de diferencias Λ' que conserva la propiedad de elipticidad para todo $p \geqslant 2$:

$$\Lambda'_y = \sum_{\alpha=1}^{p} \prod_{\beta \neq \alpha}^{1 \div p} (E + \varkappa_\beta \Lambda_\beta) \Lambda_\alpha y.$$

El operador de diferencias Λ' tiene las mismas funciones propias $\mu_k (i_1, i_2, \ldots, i_p)$ que el operador Λ construido antes, y los números

$$\lambda'_k = \sum_{\alpha=1}^{p} \prod_{\beta \neq \alpha}^{1 \div p} (1 - \varkappa_\beta \lambda_{k_\beta}^{(\beta)}) \lambda_{k_\alpha}^{(\alpha)}, \quad k = (k_1, k_2, \ldots, k_p)$$

son sus propios valores. Por lo tanto, para resolver el problema de Dirichlet de diferencias con elevado orden de exactitud y el operador Λ' indicado, se puede utilizar el algoritmo construido en el punto 1, sustituyendo λ_k por λ'_k.

3. **Tercer problema de contorno para la ecuación con derivadas mixtas en el rectángulo.** Supongamos que en el rectángulo $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \ \alpha = 1, 2\}$ hay que hallar la solución de la ecuación

$$Lu = \sum_{\alpha, \beta=1}^{2} \frac{\partial}{\partial x_\alpha} \left(k_{\alpha\beta}(x) \frac{\partial u}{\partial x_\beta} \right) - q(x) u = -\bar{f}(x),$$

$$x = (x_1, x_2) \in G, \qquad (14)$$

que en la frontera Γ satisface las condiciones de contorno

$$\begin{aligned} &k_{\alpha\alpha} \frac{\partial u}{\partial x_\alpha} + k_{\alpha\beta} \frac{\partial u}{\partial x_\beta} = \varkappa_\alpha^- u - g_\alpha^-, \quad x_\alpha = 0 \\ &-k_{\alpha\alpha} \frac{\partial u}{\partial x_\alpha} - k_{\alpha\beta} \frac{\partial u}{\partial x_\beta} = \varkappa_\alpha^+ u - g_\alpha^+, \quad x_\alpha = l_\alpha. \\ &\beta = 3 - \alpha, \ \alpha = 1, 2. \end{aligned} \qquad (15)$$

Supondremos que ha sido cumplida la condición de la elipticidad fuerte

$$c_1 \sum_{\alpha=1}^{2} \xi_\alpha^2 \leqslant \sum_{\alpha, \beta=1}^{2} k_{\alpha\beta}(x) \xi_\alpha \xi_\beta \leqslant c_2 \sum_{\alpha=1}^{2} \xi_\alpha^2, \ c_1 > 0 \qquad (16)$$

y también las condiciones

$$\varkappa_1^{\pm} = \varkappa_1^{\pm}(x_2) \geqslant 0, \quad 0 \leqslant x_2 \leqslant l_2,$$
$$\varkappa_2^{\pm} = \varkappa_2^{\pm}(x_1) \geqslant 0, \quad 0 \leqslant x_1 \leqslant l_1,$$
$$q(x) \geqslant 0. \tag{17}$$

En el rectángulo $\overline{G}$ introducimos una red rectangular arbitraria no uniforme

$$\overline{\omega} = \{x_i = (x_1(i_1), x_2(i_2)) \in \overline{G},$$
$$x_\alpha(i_\alpha) = x_\alpha(i_\alpha - 1) + h_\alpha(i_\alpha),$$
$$1 \leqslant i_\alpha \leqslant N_\alpha, \quad x_\alpha(0) = 0, \quad x_\alpha(N_\alpha) = l_\alpha, \quad \alpha = 1, 2\}$$

y determinemos el paso medio $\hbar_\alpha(i_\alpha)$:

$$\hbar_\alpha(i_\alpha) = \begin{cases} 0{,}5\, h_\alpha(1), & i_\alpha = 0, \\ 0{,}5\,[h_\alpha(i_\alpha) + h_\alpha(i_\alpha + 1)], & 1 \leqslant i_\alpha \leqslant N_\alpha - 1, \\ 0{,}5\, h_\alpha(N_\alpha), & i_\alpha = N_\alpha. \end{cases}$$

Construimos el esquema de diferencias que aproxima sobre la red $\overline{\omega}$ el problema (14), (15). Introducimos las siguientes designaciones para el flujo:

$$w_\alpha = k_{\alpha\alpha} \frac{\partial u}{\partial x} + k_{\alpha\beta} \frac{\partial u}{\partial x_\beta},$$
$$\beta = 3 - \alpha, \quad \alpha = 1, 2,$$

y anotemos el problema (14), (15) en la forma

$$Lu = \sum_{\alpha=1}^{2} \frac{\partial w_\alpha}{\partial x_\alpha} - qu = -f(x), \quad x \in G,$$
$$w_\alpha = \varkappa_\alpha^- u - g_\alpha^-, \quad x_\alpha = 0, \quad -w_\alpha = \varkappa_\alpha^+ u - g_\alpha^+$$
$$x_\alpha = l_\alpha.$$

Sobre la red $\overline{\omega}$ definimos los siguientes análogos de diferencias del flujo w_α:

$$\overset{(-)}{w_\alpha} = k_{\alpha\alpha} y_{\overline{x}_\alpha} k_{\alpha\beta} y_{\overline{x}_\beta}, \quad \overset{(+)}{w_\alpha} = k_{\alpha\alpha} y_{\overline{x}_\alpha} + k_{\alpha\beta} y_{x_\beta}$$
$$\overset{(-)}{z_\alpha} k_{\alpha\alpha} y_{x_\alpha} + k_{\alpha\beta} y_{x_\beta}, \quad \overset{(+)}{z_\alpha} = k_{\alpha\alpha} y_{x_\alpha} + k_{\alpha\beta} y_{\overline{x}_\beta}. \tag{19}$$

Empleando las designaciones introducidas, ponemos en correspondencia al problema (5) un esquema de diferencias

$$\Lambda y = (\Lambda_1 + \Lambda_2)\, y - dy = -\varphi(x), \qquad x \in \overline{\omega}, \tag{20}$$

donde $\Lambda_\alpha = \Lambda_\alpha^-$, $\alpha = 1, 2$, y el operador de diferencias Λ_α^- se determina por las fórmulas:

a) para $h_\beta(1) \leqslant x_\beta \leqslant l_\beta - h_\beta(N_\beta)$

$$\Lambda_\alpha^- y = \begin{cases} \frac{1}{2}\left[\frac{h_\beta}{\hbar_\beta}\overset{(-)}{(w_\alpha)}_{\hat{x}_\alpha} + \frac{h_\beta^{+1}}{\hbar_\beta}\overset{(-)}{(z_\alpha)}_{\check{x}_\alpha}\right], \\ h_\alpha(1) \leqslant x_\alpha \leqslant l_\alpha - h_\alpha(N_\alpha), \\ \frac{1}{2\hbar_\alpha}\left[\frac{h_\beta}{\hbar_\beta}\overset{(-)}{(w_\alpha^{+1})} + \frac{h_\beta^{+1}}{\hbar_\beta}\overset{(-)}{z_\alpha}\right] - \frac{\varkappa_\alpha^-}{\hbar_\alpha} y, \ x_\alpha = 0, \\ -\frac{1}{2\hbar_\alpha}\left[\frac{h_\beta}{\hbar_\beta}\overset{(-)}{w_\alpha} + \frac{h_\beta^{+1}}{\hbar_\beta}\overset{(-)}{z_\alpha^{-1}}\right] - \frac{\varkappa_\alpha^+}{\hbar_\alpha} y, \ x_\alpha = l_\alpha; \end{cases}$$

b) para $x_\beta = 0$

$$\Lambda_\alpha^- y = \begin{cases} \overset{(-)}{(z_\alpha)}_{\check{x}_\alpha}, \ h_\alpha(1) \leqslant x_\alpha \leqslant l_\alpha - h_\alpha(w_\alpha), \\ \frac{1}{\hbar_\alpha}\overset{(-)}{z_\alpha} - \frac{\varkappa_\alpha^-}{\hbar_\alpha} y, \ x_\alpha = 0, \\ -\frac{1}{\hbar_\alpha}\overset{(-)}{z_\alpha^{-1}} - \frac{\varkappa_\alpha^+}{\hbar_\alpha} y, \ x_\alpha = l_\alpha; \end{cases}$$

c) para $x_\beta = l_\beta$

$$\Lambda_\alpha^- y = \begin{cases} \overset{(-)}{(w_\alpha)}_{\hat{x}_\alpha}, \ h_\alpha(1) \leqslant x_\alpha \leqslant l_\alpha - h_\alpha(N_\alpha), \\ \frac{1}{\hbar_\alpha}\overset{(-)}{w_\alpha^{+1}} - \frac{\varkappa_\alpha^-}{\hbar_\alpha} y, \ x_\alpha = 0, \\ -\frac{1}{\hbar_\alpha}\overset{(-)}{w_\alpha} - \frac{\varkappa_\alpha^+}{\hbar_\alpha} y, \ x_\alpha = l_\alpha. \end{cases}$$

Aquí han sido utilizadas las designaciones

$$\overset{(-)}{(w_\alpha)}_{\hat{x}_\alpha} = \frac{\overset{(-)}{w_\alpha^{+1}} - \overset{(-)}{w_\alpha}}{\hbar_\alpha}, \quad \overset{(-)}{(z_\alpha)}_{\check{x}_\alpha} = \frac{\overset{(-)}{z_\alpha} - \overset{(-)}{z_\alpha^{-1}}}{\hbar_\alpha},$$

$$h_\beta^{+1} = h_\beta(i_\beta + 1), \ \overset{(-)}{w_\alpha^{\pm 1}} = \overset{(-)}{w_\alpha}(i_\alpha \pm 1, \ i_\beta),$$

$$\overset{(-)}{z_\alpha^{\pm 1}} = \overset{(-)}{z_\alpha}(i_\alpha \pm 1, \ i_\beta), \ d(x) = q(x).$$

El segundo miembro $\varphi(x)$ del esquema de diferencias (20) se determina según las formulas

$$\varphi(x) = \bar{f}(x) + \frac{1}{\hbar_1}\varphi_1(x) + \frac{1}{\hbar_2}\varphi_2(x),$$

donde

$$\varphi_\alpha(x)=\begin{cases} g_\alpha^-(x_\beta),\ x_\alpha=0,\\ 0,\ h_\alpha(1)\leqslant x_\alpha\leqslant l_\alpha-h_\alpha(N_\alpha),\\ g_\alpha^+(x_\beta),\ x=l_\alpha,\ \beta=3-\alpha,\ \alpha=1,\ 2 \end{cases}$$

para $0\leqslant x_\beta=l_\beta$.

En el esquema de diferencias (20) en calidad del operador Λ_α se puede emplear también el operador Λ_α^+ que se determina por las fórmulas:

a) para $h_\beta(1)\leqslant x_\beta\leqslant l_\beta-h_\beta(N_\beta)$

$$\Lambda_\alpha^+ y=\begin{cases} \frac{1}{2}\left[\frac{h_\beta^{+1}}{\hbar_\beta}\overset{(+)}{(w_\alpha)}_{\hat{x}_\alpha}+\frac{h_\beta}{\hbar_\beta}(z_\alpha^{(+)})_{\breve{x}_\alpha}\right],\ h_\alpha(1)\leqslant\\ \qquad\qquad\leqslant x_\alpha\leqslant l_\alpha-h_\alpha(N_\alpha),\\ \frac{1}{2h_\alpha}\left[\frac{h_\beta^{+1}}{\hbar_\beta}\overset{(+)}{w}_\alpha^{+1}+\frac{h_\beta}{\hbar_\beta}\overset{(+)}{z}_\alpha\right]-\frac{\varkappa_\alpha^-}{\hbar_\alpha}y,\ x_\alpha=0,\\ -\frac{1}{2h_\alpha}\left[\frac{h_\beta^{+1}}{\hbar_\beta}w_\alpha^{(+)}+\frac{h_\beta}{\hbar_\beta}\overset{(+)}{z}_\alpha^{-1}\right]-\frac{\varkappa_\alpha^+}{\hbar_\alpha}y,\ x_\alpha=l_\alpha; \end{cases}$$

b) para $x_\beta=0$

$$\Lambda_\alpha^+ y=\begin{cases} \overset{(+)}{(w_\alpha)}_{\hat{x}_\alpha},\ h_\alpha(1)\leqslant x_\alpha\leqslant l_\alpha-h_\alpha(N_\alpha),\\ \frac{1}{\hbar_\alpha}\overset{(+)}{w}_\alpha^{+1}-\frac{\varkappa_\alpha^-}{\hbar_\alpha}y,\ x_\alpha=0\\ -\frac{1}{\hbar_\alpha}\overset{(+)}{w}_\alpha-\frac{\varkappa_\alpha^+}{\hbar_\alpha}y,\ x_\alpha=l_\alpha, \end{cases}$$

c) para $x_\beta=l_\beta$

$$\Lambda_\alpha^+ y=\begin{cases} \overset{(+)}{z}_{\breve{x}_\alpha},\ h_\alpha(1)\leqslant x_\alpha\leqslant l_\alpha-h_\alpha(N_\alpha),\\ \frac{1}{\hbar_\alpha}\overset{(+)}{z}_\alpha-\frac{\varkappa_\alpha^-}{\hbar_\alpha}y,\ x_\alpha=0\\ -\frac{1}{\hbar_\alpha}\overset{(+)}{z}_\alpha^{-1}-\frac{\varkappa_\alpha^+}{\hbar_\alpha}y,\ x_\alpha=l_\alpha \end{cases}$$

Para los esquemas de diferencias examinados más arriba el molde del operador de diferencias Λ para un nodo interior de la red $\bar{\omega}$ consta de siete puntos. Si el operador de diferencias Λ_α se determina mediante la igualdad $\Lambda_\alpha=\Lambda_\alpha^0=0{,}5\,(\Lambda_\alpha^-+\Lambda_\alpha^+)$, entonces el moldo del operador Λ contendrá nueve puntos. Reducimos la forma del operador Λ_α^0:

a) para $h_\beta(1) \leqslant x_\beta \leqslant l_\beta - h_\beta(N_\beta)$

$$\Lambda^0_\alpha y = \begin{cases} (a_{\alpha\alpha} y_{\bar{x}_\alpha})_{\hat{x}_\alpha} + (k_{\alpha\beta} y^\circ_{x_\beta})^\circ_{x_\alpha}, & h_\alpha(1) \leqslant \\ \leqslant x_\alpha \leqslant l_\alpha - h_\alpha(N_\alpha), & \\ \frac{1}{h_\alpha}(a^{+1}_{\alpha\alpha} y_{x_\alpha} + k_{\alpha\beta} y^\circ_{x_\beta} - \varkappa^-_\alpha y) + & \\ + (k_{\alpha\beta} y^\circ_{x_\beta})_{x_\alpha}, \quad x_\alpha = 0, & \\ -\frac{1}{h_\alpha}(a_{\alpha\alpha} y_{\bar{x}_\alpha} + k_{\alpha\beta} y^\circ_{x_\beta} + \varkappa^+_\alpha y) + & \\ + (k_{\alpha\beta} y^\circ_{x_\beta})_{\bar{x}_\alpha}, \quad x_\alpha = l_\alpha; & \end{cases}$$

b) para $x_\beta = 0$

$$\Lambda^0_\alpha y = \begin{cases} (a_{\alpha\alpha} y_{\bar{x}_\alpha})_{\hat{x}_\alpha} + (k_{\alpha\beta} y_{x_\beta})^\circ_{x_\alpha}, & h_\alpha(1) \leqslant \\ \leqslant x_\alpha \leqslant l_\alpha - h_\alpha(N_\alpha), & \\ \frac{1}{h_\alpha}(a^{+1}_{\alpha\alpha} y_{x_\alpha} + k_{\alpha\beta} y_{x_\beta} - \varkappa^-_\alpha y) + & \\ + (k_{\alpha\beta} y_{x_\beta} y_{x_\beta})_{x_\alpha}, \quad x_\alpha = 0, & \\ -\frac{1}{h_\alpha}(a_{\alpha\alpha} y_{\bar{x}_\alpha} + k_{\alpha\beta} y_{x_\beta} + \varkappa^+_\alpha y) + & \\ + (k_{\alpha\beta} y_{x_\beta})_{\bar{x}_\alpha}, \quad x_\alpha = l_\alpha; & \end{cases}$$

c) para $x_\beta = l_\beta$

$$\Lambda^0_\alpha y = \begin{cases} (a_{\alpha\alpha} y_{\bar{x}_\alpha})_{\hat{x}_\alpha} + (k_{\alpha\beta} y_{\bar{x}_\beta})^\circ_{x_\alpha}, & h_\alpha(1) \leqslant \\ \leqslant x_\alpha \leqslant l_\alpha - h_\alpha(N_\alpha), & \\ \frac{1}{h_\alpha}(a^{+1}_{\alpha\alpha} y_{x_\alpha} + k_{\alpha\beta} y_{\bar{x}_\beta} - \varkappa^-_\alpha y) + & \\ + (k_{\alpha\beta} y_{\bar{x}_\beta})_{x_\alpha}, \quad x_\alpha = 0, & \\ -\frac{1}{h_\alpha}(a_{\alpha\alpha} y_{\bar{x}_\alpha} + k_{\alpha\beta} y_{\bar{x}_\beta} + \varkappa^+_\alpha y) + & \\ + (k_{\alpha\beta} y_{\bar{x}_\beta})_{\bar{x}_\alpha}, \quad x_\alpha = l_\alpha, & \end{cases}$$

donde $a_{\alpha\alpha} = 0{,}5\,(k_{\alpha\alpha} + k^-_{\alpha\alpha})$, $y^\circ_{x_\beta} = 0{,}5\,(y_{\hat{x}_\beta} + y_{\check{x}_\beta})$.

OBSERVACIÓN. El flujo $w_\alpha = k_{\alpha\alpha}\frac{\partial u}{\partial x_\alpha} + k_{\alpha\alpha}\frac{\partial u}{\partial x_\beta}$ se puede aproximar de otra manera, utilizando, precisamente, uno

de los análogos de diferencias:

$$w_\alpha^{(-)} = k_{\alpha\alpha}^{0,5\alpha-0,5\beta} y_{\bar{x}_\alpha} + k_{\alpha\beta}^{-0,5\alpha-0,5\beta} y_{\bar{x}_\beta},$$

$$w_\alpha^{(+)} = k_{\alpha\alpha}^{-0,5\alpha+0,5\beta} y_{\bar{x}_\alpha} + k_{\alpha\beta}^{-0,5\alpha+0,5\beta} y_{x_\beta},$$

$$z_\alpha^{(-)} = k_{\alpha\alpha}^{+0,5\alpha+0,5\beta} y_{x_\alpha} + k_{\alpha\beta}^{+0,5\alpha+0,5\beta} y_{x_\beta},$$

$$z_\alpha^{(+)} = k_{\alpha\alpha}^{+0,5\alpha-0,5\beta} y_{x_\alpha} + k_{\alpha\beta}^{+0,5\alpha-0,5\beta} y_{\bar{x}_\beta},$$

donde

$$k_{11}^{0,5_1+0,5_2} = k_{11}(x_1(i_1)+0,5h_1(i_1+1),\ x_2(i_2)+0,5h_2(i_2+1)),$$

$$k_{11}^{-0,5_1-0,5_2} = k_{11}(x_1(i_1)-0,5h_1(i_1),\ x_2(i_2)-0,5h_2(i_2))$$

etcétera. Las definiciones de los operadores Λ_α^-, Λ_α^+ y Λ_α^0 son las mismas.

Subrayemos que si la ecuación (1) no contiene las derivadas mixtas ($k_{\alpha\beta}(x) \equiv 0$), entonces los esquemas de diferencias (20), construidos con ayuda de los operadores de diferencias Λ_α^-, Λ_α^+ y Λ_α^0, diferirán unos de otros únicamente por la definición de los coeficientes de la ecuación en diferencias.

4. Métodos iterativos para resolver un problema de diferencias. El problema de contorno de diferencias (20) construido anteriormente lo anotaremos en forma de una ecuación operacional

$$Au = f \tag{21}$$

en el espacio de Hilbert H de las funciones reticulares, prefijadas sobre $\bar{\omega}$, definiendo el operador A del modo siguiente: $Ay = -\Lambda y + dy$, $y \in H$, y suponiendo que $f = \varphi$.

Definamos también el operador R que actúa en H y que es la suma de tres operadores $R = \overline{R}_1 + \overline{R}_2 + aE$, donde $a \geqslant 0$ es la constante, y $\overline{R}_1$ y $\overline{R}_2$ vienen dadas mediante las igualdades:

$$\overline{R}_\alpha y = \begin{cases} -\dfrac{1}{h_\alpha} y_{x_\alpha}, & i_\alpha = 0, \\ -y_{\bar{x}_\alpha \hat{x}_\alpha}, & 1 \leqslant i_\alpha \leqslant N_\alpha - 1, \\ \dfrac{1}{h_\alpha} y_{\bar{x}_\alpha}, & i_\alpha = N_\alpha, \end{cases} \quad \alpha = 1,\ 2.$$

Estudiemos las propiedades de los operadores A y R en el espacio H, en el cual el producto escalar lo definiremos de manera corriente:

$$(u,\ v)=\sum_{i_1=0}^{N_1}\sum_{i_2=0}^{N_2} u(i_1,\ i_2)\,v(i_1,\ i_2)\,\hbar_1(i_1)\,\hbar_2(i_2),\qquad u,\ v\in H.$$

Más abajo vamos a utilizar las designaciones siguientes:

$$(u,\ v)_\gamma=\sum_{x\in\gamma} u(x)\,v(x)\,\tau(x),$$

donde

$$\tau(x)=\tau(x_i)=\begin{cases}\hbar_\alpha(i_\alpha),\ 1\leqslant i_\alpha\leqslant N_\alpha-1,\ i_\beta=0,\ N_\beta,\ \beta\neq\alpha,\\ \hbar_\alpha(i_\alpha)+\hbar_\beta(i_\beta),\ i_\beta=0,\ N_\alpha,\quad i_\beta=0,\ N_\beta,\end{cases}$$

$$_\alpha(u,\ v)_\beta=\sum_{i_\alpha=0}^{N_\alpha=1}\sum_{i_\beta=1}^{N_\beta} uv\,h_\alpha(i_\alpha+1)\,h_\beta(i_\beta),$$

$$(u,\ v)_{\alpha\beta}=\sum_{i_\alpha=1}^{N_\alpha}\sum_{i_\beta=1}^{N_\beta} uv\,h_\alpha(i_\alpha)\,h_\beta(i_\beta),$$

$$_{\alpha\beta}(u,\ v)=\sum_{i_\alpha=0}^{N_\alpha-1}\sum_{i_\beta=0}^{N_\beta-1} uv\,h_\alpha(i_\alpha+1)\,h_\beta(i_\beta+1),$$

$$(u,\ v)_\alpha=\sum_{i_\alpha=0}^{N_\alpha}\sum_{i_\beta=0}^{N_\beta} uv\,h_\alpha(i_\alpha)\,h_\beta(i_\beta),$$

donde $\beta=3-\alpha,\quad \alpha=1,\ 2.$

Aplicando la definición del operador de diferencias $\Lambda_{\tilde\alpha}$, las fórmulas (19) y la primera fórmula de Green de diferencias, obtendremos

$$\begin{aligned}(A_{\tilde\alpha}u,\ v)&=-(\Lambda_{\tilde\alpha}u,\ v)=\\ &=-\frac{1}{2}\sum_{i_\beta=1}^{N_\beta} h_\beta(i_\beta)\Big[\sum_{i_\alpha=1}^{N_\alpha-1}\hbar_\alpha(i_\alpha)\overset{(-)}{w}{}_{\alpha\hat{x}_\alpha}v+\\ &\qquad+\overset{(-)}{w}{}_\alpha^{i_\alpha+1}v\big|_{i_\alpha=0}-\overset{(-)}{w}{}_\alpha v\big|_{i_\alpha=N_\alpha}\Big]-\\ &-\frac{1}{2}\sum_{i_\beta=0}^{N_\beta-1} h_\beta(i_\beta+1)\Big[\sum_{i_\alpha=1}^{N_\alpha-1}\hbar_\alpha(i_\alpha)\overset{(-)}{z}{}_{\alpha\overset{\vee}{x}_\alpha}v+\\ &\qquad+\overset{(-)}{z}{}_\alpha v\big|_{i_\alpha=0}-\overset{(-)}{z}{}_\alpha^{i_\alpha+1}v\big|_{i_\alpha=N_\alpha}\Big]+\end{aligned}$$

$$+\sum_{i_\beta=0}^{N_\beta} \hbar_\beta(i_\beta)\,[\varkappa_\alpha^- uv|_{i_\alpha=0}+\varkappa_\alpha^+ uv|_{i_\alpha=N_\alpha}]=$$

$$=\frac{1}{2}\sum_{i_\beta=1}^{N_\beta}\sum_{i_\alpha=1}^{N_\alpha} h_\alpha(i_\alpha)\,h_\beta(i_\beta)\overset{(-)}{w}_\alpha v_{\bar{x}_\alpha}+$$

$$+\frac{1}{2}\sum_{i_\beta=0}^{N_\beta-1}\sum_{i_\alpha=0}^{N_\alpha-1} h_\alpha(i_\alpha+1)\,h_\beta(i_\beta+1)\overset{(-)}{z}_\alpha v_{x_\alpha}+$$

$$+\sum_{i_\beta=0}^{N_\beta}\hbar_\beta(i_\beta)\,[\varkappa_\alpha^- uv|_{i_\alpha=0}+\varkappa_\alpha^+ uv|_{i_\alpha=N_\alpha}]=$$

$$=0{,}5\,[(k_{\alpha\alpha}u_{\bar{x}_\alpha}+k_{\alpha\beta}u_{\bar{x}_\beta},\ v_{\bar{x}_\alpha})_{\alpha\beta}+$$

$$+{}_{\alpha\beta}(k_{\alpha\alpha}u_{x_\alpha}+k_{\alpha\beta}u_{x_\beta},v_{x_\beta})]+\sum_{i_\beta=0}^{N_\beta}\hbar_\beta(i_\beta)\,[\varkappa_\alpha^- uv|_{i_\alpha=0}+$$

$$+\varkappa_\alpha^+ uv|_{i_\alpha=N_\alpha}],\ \beta=3-\alpha,\ \alpha=1,\ 2.$$

De aquí para el operador A hallaremos

$$(Au,\ v)=\sum_{\alpha=1}^{2}(A_\alpha^- u,\ v)+(du,\ v)=0{,}5\,[(k_{11}u_{\bar{x}_1}v_{\bar{x}_1}+$$

$$+k_{12}u_{\bar{x}_2}v_{\bar{x}_1}+k_{21}u_{\bar{x}_1}v_{\bar{x}_2}+k_{22}u_{\bar{x}_2}v_{\bar{x}_2},\ 1)_{12}+$$

$$+{}_{12}(k_{11}u_{x_1}v_{x_1}+k_{12}u_{x_2}v_{x_1}+k_{21}u_{x_1}v_{x_2}+$$

$$+k_{22}u_{x_2}v_{x_2},\ 1)]+(\varkappa u,\ v)_\gamma+(du,\ v=$$

$$=0{,}5\,[(\sum_{\alpha,\beta=1}^{2}k_{\alpha\beta}u_{x_\beta}v_{x_\alpha},\ 1)_{12}+$$

$$+{}_{12}(\sum_{\alpha,\beta=0}^{2}k_{\alpha\beta}u_{x_\beta}v_{x_\alpha},\ 1)]+(\varkappa u,\ v)_\gamma+(du,\ v). \quad (22)$$

Aquí la función $\varkappa$ está definida sobre γ de la manera siguiente:

$$\varkappa(x_i)=\begin{cases}\varkappa_\alpha^-,\ i_\alpha=0,\ 1_\beta\leqslant i_\beta\leqslant N_\beta-1,\\ \varkappa_\alpha^+,\ i_\alpha=N_\alpha,\ 1\leqslant i_\beta\leqslant N_\beta-1,\\ \dfrac{\varkappa^\pm\hbar_\beta+\varkappa_\beta^\pm h_\alpha}{\hbar_\beta+h_\alpha},\ i_\alpha=0,\ N_\alpha\ y\ i_\beta=0,\ N_\beta.\end{cases}$$

De (22) se deduce que el operador A es autoconjugado en H, si está cumplida la condición

$$k_{12} = k_{21}. \tag{23}$$

Suponiendo en (22) $v = u$ y empleando las condiciones (16), obtendremos la desigualdad

$$(Au,\ u) \geqslant 0{,}5c_1 \sum_{\alpha=1}^{2} [(u^2_{\bar{x}_\alpha},\ 1)_{12} + {}_{12}(u^2_{\bar{x}_\alpha},\ 1)] +$$

$$+ (du^2,\ 1) + (\varkappa u^2,\ 1)_\gamma = c_1 \sum_{\alpha=1}^{2} (u^2_{\bar{x}_\alpha},\ 1)_\alpha +$$

$$+ (du^2,\ 1) + (\varkappa u^2,\ 1)_\gamma \tag{24}$$

y otra desigualdad analógica

$$(Au,\ u) \leqslant c_2 \sum_{\alpha=1}^{2} (u^2_{\bar{x}_\alpha},\ 1)_\alpha + (du^2,\ 1) + (\varkappa u^2,\ 1)_\gamma. \tag{25}$$

Si en el esquema (20) en calidad de Λ_α se utiliza el operador Λ^+_α, entonces para el operador respectivo A en lugar de (22) tendremos la identidad

$$(Au,\ v) = 0{,}5\,[{}_2(k_{11}u_{\bar{x}_1}v_{\bar{x}_1} + k_{12}u_{x_1}v_{\bar{x}_1} + k_{21}u_{\bar{x}_1}v_{x_2} +$$

$$+ k_{22}u_{x_2}v_{x_2},\ 1)_1 + {}_1(k_{11}u_{x_1}v_{x_1} + k_{12}u_{\bar{x}_2}v_{x_1} +$$

$$+ k_{21}u_{x_1}v_{\bar{x}_2} + k_{22}u_{\bar{x}_2}v_{\bar{x}_2},\ 1)_2] + (du,\ v) + (\varkappa u,\ v)_\gamma,$$

del cual se desprende la autoconjugación del operador A al satisfacer la condición (23). Las desigualdades (24), (25) serán válidas. Dichas afirmaciones quedarán justas también para el caso, cuando en el esquema (20) como Λ_α se utiliza el operador Λ^0_α.

La autoconjugación del operador R en H y las igualdades

$$(Ru,\ u) = \sum_{\alpha=1}^{2} (u^2_{\bar{x}_\alpha},\ 1)_\alpha + a\,(u^2,\ 1), \tag{26}$$

para toda $u \in H$ se deducen de las fórmulas de Green de diferencias.

Para nosotros el problema consiste en encontrar las constantes m_1 y m_2 en las desigualdades

$$m_1\,(Ru,\ u) \leqslant (Au,\ u) \leqslant m_2\,(Ru,\ u), \tag{27}$$

es decir, las constantes de la equivaloncia energética de los operadores A y R. Resolveremos dicho problema, partiendo

de que estén cumplidas las condiciones (17) y los coeficientes $q(x)$ y $\varkappa_\alpha^+(x_\beta)$ no sean iguales idénticamente a cero.

A fin de hallar m_1 y m_2 nos harán falta dos lemas [1]).

LEMA 1. *Sea* ω^* *cierto conjunto de nodos de la red* $\overline{\omega}$, *además*

$$\sum_{x\in\omega^*} \hbar_1(x_1)\,\hbar_2(x_2) = s > 0.$$

Para cualquier función reticular $u(x)$ *prefijada en* $\overline{\omega}$, *tiene lugar la desigualdad*

$$\nu \sum_{\alpha=1}^{2} (u_{\bar{x}_\alpha}^2,\ 1)_\alpha + \mu\,(u^2,\ 1)_{\omega^*} \geqslant \mu M\,(u^2,\ 1), \qquad (28)$$

donde μ *es una constante arbitraria,* ν, *la constante que satisface la condición* $\nu > \max\,[0,\ -0{,}5\mu\,(l_1^2 + l_2^2)]$, *y*

$$M = \frac{2s\nu}{l_1 l_2\,[2\nu + \mu\,(l_1^2 + l_2^2)]} > 0.$$

Expondremos las principales etapas de demostración del lema, remitiendo al libro indicado en la cita si el lector quiere entrar en detalles. Supongamos que $x = (x_1,\ x_2)$ y $y = (y_1,\ y_2)$ son dos nodos arbitrarios de la red $\overline{\omega}$. Para toda función reticular u dada sobre $\overline{\omega}$, si $\varepsilon_1 > 0$ y $\varepsilon_2 > 0$ es cierta la desigualdad

$$(1-\varepsilon_1)\,u^2(x) \leqslant \left(\frac{1}{\varepsilon_1} - 1\right) u^2(y) + (1+\varepsilon_2)\,|x_1 - y_1| \sum_{x_1\in\omega_1^+} u_{\bar{x}_1}^2(x_1,\ x_2)\,h_1(x_1) + \left(1 + \frac{1}{\varepsilon_2}\right)|x_2 - y_2| \sum_{x_2\in\omega_2^+} u_{\bar{x}_2}^2(y_1,\ x_2)\,h_2(x_2).$$

Multipliquemos esta desigualdad por $\hbar_1(x_1)\,\hbar_2(x_2)$, sumemos respecto a $x \in \overline{\omega}$ y tomemos en consideración la desigualdad

$$\sum_{x_\alpha\in\overline{\omega}_\alpha} |x_\alpha - y_\alpha|\,\hbar_\alpha(x_\alpha) \leqslant 0{,}5 l_\alpha^2,\ \ \alpha = 1,\ 2.$$

En resumen obtendremos

$$(1-\varepsilon_1)\,(u^2,\ 1) \leqslant l_1 l_2 \left(\frac{1}{\varepsilon_1} - 1\right) u^2(y) + 0{,}5 l_1^2\,(1+\varepsilon_2)\,(u_{\bar{x}_1}^2,\ 1)_1 + 0{,}5 l_1 l_2^2 \left(1 + \frac{1}{\varepsilon_2}\right) \sum_{x_\alpha\in\omega_2^+} u_{\bar{x}_\alpha}^2(y_1,\ x_2)\,h_2(x_2).$$

[1]) Véanse los lemas 9 y 10 del § 5 del cap. V del libro de A. A. Samarski, V. B. Andreev, Métodos de diferencias para las ecuaciones elípticas, Editorial MIR, 1979, Moscú.

Multipliquemos esta desigualdad por $h_1(y_1)\,h_2(y_2)$ y sumemos respecto a $y \in \omega^*$. Teniendo en cuenta la estimación $s \leqslant l_1 l_2$, obtendremos

$$s(1-\varepsilon_1)(u^2, 1) \leqslant l_1 l_2 \left(\frac{1}{\varepsilon_1} - 1\right)(u^2, 1)_{\omega^*} + \\ + 0{,}51_1^3 l_2 (1+\varepsilon_2)\, u^2_{\bar{x}_1}, 1)_1 + \\ + 0{,}51_1 l_2^3 \left(1 + \frac{1}{\varepsilon_2}\right)(u^2_{\bar{x}_2}, 1)_2. \qquad (29)$$

Poniendo aquí

$$\varepsilon_1 = \frac{2\nu}{2\nu + \mu(l_1^2 + l_2^2)}, \quad \varepsilon_2 = \left(\frac{l_2}{l_1}\right)^2,$$

obtenemos la estimación (28). El lema queda demostrado.

LEMA. 2. *Sea* γ^* *ciero conjunto de nodos de la frontera* γ, *además* $\sum\limits_{x \in \gamma^*} \tau(x) \geqslant \delta$.

Para cualquera función reticular $u(x)$ *prefijada sobre* $\bar{\omega}$, *tiene lugar la desigualdad*

$$\nu \sum_{\alpha=1}^{2} (u^2_{\bar{x}_\alpha}, 1)_\alpha + \mu (u^2, 1)_{\gamma^*} \geqslant \mu M (u^2, 1), \qquad (30)$$

donde μ *es una constante arbitraria*; ν, *constante que satisfase la condición* $\nu > \max[0, -\mu(l_1^2 + l_2^2)(1/l_1 + 1/l_2)]$, y

$$M = \frac{\sigma\nu}{\nu l_1 l_2 + \mu(l_1^2 + l_2^2)(l_1 + l_2)} > 0.$$

En la demostracion del lema 2 juega el papel principal la desigualdad análoga a (29)

$$\sigma(1-\varepsilon_1)(u^2, 1) \leqslant l_1 l_2 \left(\frac{1}{\varepsilon_1} - 1\right)(u^2, 1)_{\gamma^*} + l_1^2 \times$$

$$\times (l_1 + l_2)(1+\varepsilon_2)(u^2_{\bar{x}_1}, 1) + l_2^2 (l_1 + l_2) \left(1 + \frac{1}{\varepsilon_2}\right)(u^2_{\bar{x}_2}, 1)_2.$$

Poniendo aquí

$$\varepsilon_1 = \frac{\nu l_1 l_2}{\nu l_1 l_2 + \mu(l_1^2 + l_2^2)(l_1 + l_2)}, \quad \varepsilon_2 = \left(\frac{l_2}{l_1}\right)^2,$$

obtendremos la estimación (30).

Ahora aprovechemos los lemas 1 y 2 para estimar m_1 y m_2 en las desigualdades (27). Designemos con ω^* los nodos de la red $\bar{\omega}$, en los cuales $d(x) \neq 0$; con γ^*, los nodos de la

frontera γ, en los cuales $\varkappa(x) \not\equiv 0$ y adoptemos que

$$d_* = \min_{x \in \omega^*} d(x), \quad d^* = \max_{x \in \omega^*} d(x),$$

$$\varkappa_* = \min_{x \in \gamma^*} \varkappa(x), \quad \varkappa^* = \max_{x \in \gamma^*} \varkappa(x).$$

Poniendo en (28) $\mu = d_*$, $\nu = \nu_1 > 0$ y en (30) $\mu = \varkappa_*$, $\nu = \nu_2 > 0$, de (24) obtendremos la estimación

$$(Au, u) \geqslant c_1 \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha + (du^2, 1)_{\omega^*} +$$

$$+ (\varkappa u^2, 1)_{\gamma^*} \geqslant (c_1 - \nu_1 - \nu_2) \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha +$$

$$+ \nu_1 \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha + d_* (u^2, 1)_{\omega^*} + \nu_2 \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha +$$

$$+ \varkappa_* (u^2, 1)_{\gamma^*} \geqslant (c_1 - \gamma_1 - \gamma_2) \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha + M_1 (u^2, 1), \tag{31}$$

donde

$$M_1 = \frac{2s\nu_1 d_*}{l_1 l_2 [2\nu_1 + d_* (l_1^2 + l_2^2)]} + \frac{\sigma \nu_2 \varkappa_*}{\nu_2 l_1 l_2 + \varkappa_* (l_1^2 + l_2^2)(l_1 + l_2)}.$$

A continuación, suponiendo en (28) $\mu = -d^*$, $\nu = \nu_3 > 0{,}5 d^* (l_1^2 + l_2^2)$ y en (30) $\mu = -\varkappa^*$, $\nu = \nu_4 > \varkappa^* (l_1^2 + l_2^2) \times (1/l_2 + 1/l_2)$, de (25) obtendremos la estimación

$$(Au, u) \leqslant c_2 \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha + (du^2, 1)_{\omega^*} +$$

$$+ (\varkappa u^2, 1)_{\gamma^*} \leqslant (c_2 + \nu_3 + \nu_4) \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha -$$

$$- [\nu_3 \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha - d^* (u^2, 1)_{\omega^*}] - [\nu_4 \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha -$$

$$- \varkappa^* (u^2, 1)_{\gamma^*}] \leqslant (c_2 + \nu_3 + \nu_4) \sum_{\alpha=1}^{2} (u_{x_\alpha}^2, 1)_\alpha + M_2 (u^2, 1), \tag{32}$$

donde

$$M_2 = \frac{2s\nu_3 d^*}{l_1 l_2 [2\nu_3 - d^* (l_1^2 + l_2^2)[} + \frac{\sigma \nu_4 \varkappa^*}{\nu_4 l_1 l_2 - \varkappa^* (l_1^2 + l_2^2)(l_1 + l_2)}.$$

Sea a la constante que entra en la definición del operador R. Eligimos los parámetros ν_2 y ν_4 de tal modo que sean válidas las igualdades

$$\frac{M_1}{c_1-\nu_1-\nu_2}=a, \quad \frac{M_2}{c_2+\nu_3+\nu_4}=a, \tag{33}$$

donde M_1 y M_2 fueron determinadas más arriba. De (33) obtendremos que $\nu_2=\nu_2(a, \nu_1)$ y $\nu_4=\nu_4(a, \nu_3)$. Notemos que la región de los valores admisibles de ν_1 y ν_3 se determina mediante las restricciones:

$$\nu_1>0, \quad \nu_2>0, \quad c_1-\nu_1-\nu_2>0,$$
$$\nu_3>0{,}5d^*(l_1^2+l_2^2), \quad \nu_4>\varkappa^*(l_1^2+l_2^2)(1/l_1+1/l_2).$$

En las regiones indicadas tomemos los valores de ν_1 y ν_3 partiendo de las condiciones del máximo de $c_1-\nu_1-\nu_2(a, \nu_1)$ respecto a la variable ν_1 y del mínimo de $c_2+\nu_3+\nu_4(a, \nu_3)$ respecto a la variable ν_3. Designemos

$$m_1=m_1(a)=\max_{\nu_1}[c_1-\nu_1-\nu_2(a, \nu_1)],$$
$$m_2=m_2(a)=\min_{\nu_3}[c_2+\nu_3+\nu_4(a, \nu_3)]. \tag{34}$$

De (31), (32) y (26) obtendremos que son ciertas las desigualdades (27) con los valores de m_1 y m_2 indicados en (34). Queda por elegir el valor óptimo de la constante a, partiendo de la condición del máximo de la razón $m_1(a)/m_2(a)$.

Notemos que m_1 y m_2 no dependen de los pasos de la red $\bar{\omega}$, si s y σ son acotados por abajo por magnitudes que no dependan de la red $\bar{\omega}$.

El operador elegido R lo utilizaremos como un regularizador al construir los esquemas iterativos para resolver la ecuación (21). Limitarémonos a examinar los esquemas de dos capas implícitos

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f, \quad k=0, 1, \ldots, \quad y_0\in H \tag{35}$$

con el operador B autoconjugado y con el conjunto de Chebishev de parámetros $\{\tau_k\}$. Recordemos que para el esquema (35) deben ser dadas las constantes γ_1 y γ_2 en las desigualdades

$$\gamma_1 B\leqslant A\leqslant\gamma_2 B, \quad \gamma_1>0, \tag{36}$$

y el número de iteraciones n se estima mediante la magnitud $n_0(\varepsilon)$, donde

$$n_0(\varepsilon) = \frac{\ln 0{,}5\varepsilon}{\ln \rho_1}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Si la red $\bar{\omega}$ es uniforme en una de las direcciones, entonces adoptemos $B = R$. En este caso en (36) $\gamma_1 = m_1$ y $\gamma_2 = m_2$ y la ecuación $Ry_{k+1} = (R - \tau_{k+1}A)\, y_k + \tau_{k+1} f$ con el segundo miembro conocido podrá ser resuelta ya sea por el método de reducción total (véase el cap. III), ó mediante el método de separación de variables (véase el cap. IV) gastando $O(N^2 \log_2 N)$ operaciones aritméticas ($N_1 = N_2 = N = 2^n$). Puesto que m_1 y m_2 no dependen de la cantidad de nodos en la red $\bar{\omega}$, el número total de operaciones gastadas para resolver aproximadamente con exactitud ε la ecuación (21) se estima por la magnitud $Q(\varepsilon) = O(N^2 \log_2 N \mid \ln 0{,}5\varepsilon \mid)$.

Si la red $\bar{\omega}$ es no uniforme respecto a cada dirección, entonces ponemos $B = (E + \omega_1 R_1)(E + \omega_2 R_2)$, donde $R_\alpha = \bar{R}_\alpha + 0{,}5aE$, $\alpha = 1, 2$, y $\bar{R}_\alpha$ han sido determinados más arriba.

Está claro que los operadores R_1 y R_2 son autoconjugados en H y conmutables. Además, en virtud del carácter no negativo del operador $\bar{R}_\alpha$ tenemos la estimación

$$R_\alpha \geqslant \delta_\alpha E, \quad \alpha = 1, 2, \quad \delta_1 = \delta_2 = 0{,}5a.$$

Del lema 17 del cap. V encontramos la estimación

$$R_\alpha \leqslant \Delta_\alpha E, \quad \Delta_\alpha = \max\left[\frac{4}{h_\alpha(1)}, \frac{4}{h_\alpha(N_\alpha)}, \max_{1 \leqslant i_\alpha \leqslant N_\alpha - 1} \frac{4}{h_\alpha(i_\alpha)\, h_\alpha(i_\alpha + 1)}\right].$$

Aplicando el teorema 3 del capítulo presente obtenemos los valores óptimos de los parámetros ω_1, ω_2, así como también las constantes de la equivalencia energética $\overset{\circ}{\gamma}_1$ y $\overset{\circ}{\gamma}_2$ de los operadores B y R:

$$\overset{\circ}{\gamma}_1 B \leqslant R \leqslant \overset{\circ}{\gamma}_2 B, \quad \overset{\circ}{\gamma}_1 > 0.$$

De aquí y de (27) hallamos que en (36) $\gamma_1 = m_1 \overset{\circ}{\gamma}_1$ y $\gamma_2 = m_2 \overset{\circ}{\gamma}_2$.

Se puede mostrar que para el método que se examina $\gamma_1 = O(1)$ y $\gamma_2 = O(1/h)$, donde h es el paso mínimo de la

red $\overline{\omega}$. Por lo tanto, el número de iteraciones será estimado por la magnitud $O(h^{-1/2} \mid \ln 0{,}5\varepsilon \mid)$. El número de operaciones aritméticas gastadas para una iteración es proporcional al número de incognitas en nuestro problema, ya que para la inversión de los operadores $E + \omega_\alpha R_\alpha$, $\alpha = 1, 2$, se puede emplear el método de factorización.

De esta manera, utilizando el principio de regularización, hemos construido dos métodos iterativos, por medio de los cuales es posible encontrar la solución aproximada del análogo de diferencias del tercer problema de contorno para la ecuación con derivadas mixtas en un rectángulo.

§ 3. Sistemas de ecuaciones elípticas

1. Problema de Dirichlet para un sistema de ecuaciones elípticas en un paralelepípedo p-dimensional. Sean $u = (u^1(x), u^2(x), \ldots, u^{m_0}(x))$ y $f = (f^1(x), f^2(x), \ldots, f^{m_0}(x))$ vectores de dimensión m_0, $x = (x_1, x_2, \ldots, x_p)$ un punto del espacio de p-dimensiones, y $k = (k_{\alpha\beta})$ es la matriz celular de dimensión $p \times p$, tal que la célula $k_{\alpha\beta} = (k_{\alpha\beta}^{sm}(x))$ es la matriz de dimensión $m_0 \times m_0$:

$$k = \begin{Vmatrix} k_{11} & k_{12} & \ldots & k_{1p} \\ k_{21} & k_{22} & \ldots & k_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ k_{p1} & k_{p2} & \ldots & k_{pp} \end{Vmatrix},$$

$$k_{\alpha\beta} = \begin{Vmatrix} k_{\alpha\beta}^{11} & k_{\alpha\beta}^{12} & \ldots & k_{\alpha\beta}^{1m_0} \\ k_{\alpha\beta}^{21} & k_{\alpha\beta}^{22} & \ldots & k_{\alpha\beta}^{2m_0} \\ \ldots & \ldots & \ldots & \ldots \\ k_{\alpha\beta}^{m_0 1} & k_{\alpha\beta}^{m_0 2} & \ldots & k_{\alpha\beta}^{m_0 m_0} \end{Vmatrix}.$$

En el paralelepípedo p-dimensional $\overline{G} = \{0 \leqslant x_\alpha \leqslant l_\alpha, \alpha = 1, 2, \ldots, p\}$ con frontera Γ examinaremos el problema de Dirichlet para el sistema de ecuaciones elípticas:

$$Lu = \sum_{\alpha, \beta=1}^{p} \frac{\partial}{\partial x_\alpha}\left(k_{\alpha\beta} \frac{\partial u}{\partial x_\beta}\right) = -f(x), \quad x \in G,$$

$$u(x) = g(x) \quad x \in \Gamma. \qquad (1)$$

Si pasamos de la escritura vectorial a la escalar, entonces el problema (1) se escribe en forma del sistema

$$(Lu)^s = -f^s(x), \quad x \in G,$$
$$u^s(x) = g^s(x), \quad x \in \Gamma, \quad s = 1, 2, \ldots, m_0,$$

donde

$$(Lu)^s = \sum_{\alpha, \beta=1}^{p} \sum_{m=1}^{m_0} \frac{\partial}{\partial x_\alpha} \left(k_{\alpha\beta}^{sm}(x) \frac{\partial u^m}{\partial x_\beta} \right). \tag{2}$$

Supondremos, que está cumplida la condición de elipticidad fuerte

$$c_1 \sum_{\alpha=1}^{p} |\xi_\alpha|^2 \leqslant \sum_{\alpha, \beta=1}^{p} (k_{\alpha\beta}\xi_\alpha, \xi_\beta) \leqslant c_2 \sum_{\alpha=1}^{p} |\xi_\alpha|^2, \tag{3}$$

donde $c_1 > 0$ y $c_2 > 0$ son constantes, que no dependen de x, $\xi_\alpha = (\xi_\alpha^1, \xi_\alpha^2, \ldots, \xi_\alpha^{m_0})$, $\alpha = 1, 2, \ldots, p$ son los vectores arbitrarios,

$$|\xi_\alpha|^2 = \sum_{s=1}^{m_0} (\xi_\alpha^s)^2, \quad (k_{\alpha\beta}\xi_\alpha, \xi_\beta) = \sum_{s, m=1}^{m_0} k_{\alpha\beta}^{sm} \xi_\alpha^s \xi_\beta^m.$$

Señalemos, que la desigualdad izquierda en (3) significa la definición positiva de la matriz k.

Construyamos un esquema de diferencias que aproxima el problema (1). Para esto introduzcamos en la región $\overline{G}$ una red rectangular uniforme

$$\overline{\omega} = \{x_i = (i_1 h_1, \ldots, i_p h_p) \in \overline{G}, \quad 0 \leqslant i_\alpha \leqslant N_\alpha, \quad h_\alpha N_\alpha = l_\alpha, \quad \alpha = 1, 2, \ldots, p\}$$

con frontera γ, de manera que $\overline{\omega} = \omega \cup \gamma$. Sobre la red $\overline{\omega}$ examinaremos las funciones reticulares vectoriales, cuyas componentes son las funciones reticulares de p variables discretas, por ejemplo $y = (y^1, y^2, \ldots, y^{m_0})$, y al mismo tiempo $y^s = y^s(x_i) = y^s(i_1, i_2, \ldots, i_p)$.

El problema de Dirichlet de diferencias para el sistema (1) sobre la red $\overline{\omega}$ en la escritura vectorial tiene la forma

$$\Lambda y = \sum_{\alpha, \beta=1}^{p} 0{,}5 \left[(k_{\alpha\beta} y_{x_\beta})_{x_\alpha} + (k_{\alpha\beta} y_{x_\beta})_{\overline{x}_\alpha} \right] = -\varphi(x), \quad x \in \omega,$$
$$y(x) = g(x), \quad x \in \gamma$$

Pasando a la escritura escalar, obtendremos el sistema

$$\begin{aligned} &(\Lambda^- y)^s = -\varphi^s(x), \quad x \in \omega \\ &y^s(x) = g^s(x), \quad x \in \gamma, \quad s = 1, 2, \ldots, m_0, \end{aligned} \tag{4}$$

donde

$$(\Lambda^- y)^s = \sum_{\alpha,\beta=1}^{p} \sum_{m=1}^{m_0} 0{,}5\,[(k^{sm}_{\alpha\beta} y^m_{\bar{x}_\alpha})_{x_\alpha} + (k^{sm}_{\alpha\beta} y^m_{x_\beta})_{\bar{x}_\alpha}].$$

El operador Λ^-, al igual que en el caso de una ecuación elíptica escalar, admite otra escritura, precisamente:

$$\Lambda^- y = \sum_{\alpha=1}^{p} 0{,}5\,[(k_{\alpha\alpha} y_{\bar{x}_\alpha})_{x_\alpha} + (k_{\alpha\alpha} y_{x_\alpha})_{\bar{x}_\alpha}] + \sum_{\alpha\neq\beta}^{1\div p} 0{,}5\,[(k_{\alpha\beta} y_{\bar{x}_\beta})_{x_\alpha} + (k_{\alpha\beta} y_{x_\beta})_{\bar{x}_\alpha}].$$

Señalemos, que para aproximar el operador diferencial L se pueden también utilizar otros operadores diferentes de Λ^-, por ejemplo

$$\Lambda^+ y = \sum_{\alpha=1}^{p} 0{,}5\,[(k_{\alpha\alpha} y_{\bar{x}_\alpha})_{x_\alpha} + (k_{\alpha\alpha} y_{x_\alpha})_{\bar{x}_\alpha}] + \sum_{\alpha\neq\beta}^{1\div p} 0{,}5\,[(k_{\alpha\beta} y_{x_\beta})_{x_\alpha} + (k_{\alpha\beta} y_{\bar{x}_\beta})_{\bar{x}_\alpha}]$$

ó

$$\Lambda^0 y = 0{,}5\,(\Lambda^- + \Lambda^+)\,y = \sum_{\alpha=1}^{p} 0{,}5\,[(k_{\alpha\alpha} y_{\bar{x}_\alpha})_{x_\alpha} + (k_{\alpha\alpha} y_{x_\alpha})_{\bar{x}_\alpha}] + \sum_{\alpha\neq\beta}^{1\div p} (k_{\alpha\beta} y^{\circ}_{x_\beta})^{\circ}_{x_\alpha}.$$

Introduzcamos el espacio H de las funciones vectoriales reticulares, prefijada sobre ω, y definamos en él el producto escalar

$$(u, v) = \sum_{s=1}^{m_0} (u^s, v^s), \quad (u^s, v^s) = \sum_{x\in\omega} u^s(x)\, v^s(x)\, h_1 h_2 \ldots h_p),$$

$$u = (u^1, u^2, \ldots, u^{m_0}), \quad v = (v^1, v^2, \ldots, v^{m_0}), \quad u, v \in H.$$

Definiremos el operador de Laplace de diferencias:

$$\mathcal{R}y = \sum_{\alpha=1}^{p} y_{\bar{x}_\alpha x_\alpha}, \quad (\mathcal{R}y)^s = \sum_{\alpha=1}^{p} y^s_{\bar{x}_\alpha x_\alpha}.$$

En el espacio H definames los operadores A y R de la manera usual:

$$Ay = -\Lambda \overset{\circ}{y}, \quad Ry = -\mathscr{P}\overset{\circ}{y}, \quad y \in H,$$

donde $\overset{\circ}{y}(x) = y(x)$ para $x \in \omega$ y $\overset{\circ}{y}(x) = 0$, si $x \in \gamma$.

Utilizando las notaciones introducidas y corrigiendo de manera obvia el miembro derecho φ de la ecuación (4) en los nodos fronterizos, escribamos el problema de diferencias (4) en la forma de la ecuación operacional

$$Au = f, \tag{5}$$

definida en el espacio de Hilbert H.

Utilizando la fórmula de Green en diferencias para funciones reticulares escalares y las condiciones (3) y suponiendo que se cumplen las condiciones de simetría

$$k_{\alpha\beta}^{sm} = k_{\beta\alpha}^{ms}, \quad \alpha, \beta = 1, 2, \ldots, p, \quad s, m = 1, 2, \ldots, m_0, \tag{6}$$

obtendremos que los operadores R y A son autoconjugados en H y energéticamente equivalentes con las constantes c_1 y c_2, es decir, tienen lugar las desigualdades operacionales

$$c_1 R \leqslant A \leqslant c_2 R, \quad c_1 > 0. \tag{7}$$

Para encontrar una solución aproximada de la ecuación (5) aprovecharemos el método iterativo implícito de dos capas con parámetros de Chebishev

$$B\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, y_0 \in H, \tag{8}$$

donde

$$\tau_k = \frac{\tau_0}{1 + \rho_0 \mu_k}, \quad \mu_k \in \mathfrak{M}_n = \left\{ -\cos\frac{(2i-1)\pi}{2n}, \quad 1 \leqslant i \leqslant n \right\},$$

$$k = 1, 2, \ldots, n,$$

$$\tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_3},$$

$$n \geqslant n_0(\varepsilon) = \ln(0{,}5\varepsilon)/\ln\rho_1,$$

y γ_1, γ_2 son las constantes de equivalencias energéticas de los operadores autoconjugados A y B:

$$\gamma_1 B \leqslant A \leqslant \gamma_2 B, \quad \gamma_1 > 0, \quad A = A^*, \quad B = B^*. \tag{9}$$

Si en calidad del operador B elegimos el operador R definido más arriba, entonces de (7) obtendremos, que en las

desigualdades (9) $\gamma_1 = c_1$ y $\gamma_2 = c_2$. Por lo tanto, el número de iteraciones del método (8) no depende en el caso examinado del número de nodos de la red: $n = O(\ln(2/\varepsilon))$.

De la definición de los operadores A y B se deduce que para encontrar y_{k+1} por la aproximación anterior y_k ya conocida es necesario resolver el siguiente problema en diferencias:

$$\mathscr{R} y_{k+1} = -F_k, \quad x \in \omega, \quad F_k = \tau_{k+1}[(\Lambda^- y_k + \varphi) - \mathscr{R} y_k],$$
$$y_{k+1} = g, \quad x \in \gamma.$$

En forma escalar este problema se escribe en la forma del sistema

$$\sum_{\alpha=1}^{p} (y^s_{k+1})_{\bar{x}_\alpha x_\alpha} = -F^s_k(x), \quad x \in \omega,$$
$$y^s_{k+1}(x) = g^s(x), \quad x \in \gamma, \quad s = 1, 2, \ldots, m_0. \tag{10}$$

Ya que cada ecuación del sistema (10) puede ser resuelta independientemente de las otras ecuaciones, entonces la búsqueda de la aproximación y_{k+1} se reduce a la resolución de m_0 problemas de Dirichlet de diferencias en un paralelepípedo p-dimensional sobre la red rectangular uniforme $\overline{\omega}$.

Si para resolver el problema de Dirichlet p-dimensional de diferencias para la ecuación de Poisson utilizamos el método de separación de variables con el algoritmo de la transformación de Fourier discreta rápida, entonces se puede mostrar, que se exige $q \approx 4pN^p \log_2 N$ ($N_1 = N_2 = \ldots = N_p = N = 2^{n_0}$) operaciones aritméticas. Por consiguiente, para la resolución del sistema (10) se exige $Q_{m_0} = m_0 q$ operaciones, y en total para encontrar la solución del problema en diferencias (4) con una exactitud ε es necesario gastar $Q = nQ_{m_0} = nm_0 q = O(m_0 pN^p) \ln \frac{2}{\varepsilon} \log_2 N)$ operaciones aritméticas.

Examinemos ahora el método iterativo alternativo triangular. El esquema iterativo tiene la forma (8), donde B es el operador factorizado $B = (E + \omega R_1)(E + \omega R_2)$, $R_1 = R_2^*$, $R_1 + R_2 = R$. Los operadores R_1 y R_2 se definen con ayuda de los operadores de diferencias $\mathscr{R}_1$ y $\mathscr{R}_2$ de la forma siguiente: $R_\alpha y = -\mathscr{R}_\alpha y$, $\alpha = 1, 2$, $y(x) = \overset{\circ}{y}(x)$ para $x \in \omega$ y $\overset{\circ}{y}(x) = 0$ para $x \in \gamma$, donde

$$\mathscr{R}_1 y = -\sum_{\alpha=1}^{p} \frac{1}{h_\alpha} y_{\bar{x}_\alpha}, \quad \mathscr{R}_2 y = \sum_{\alpha=1}^{p} \frac{1}{h_\alpha} y_{x_\alpha}.$$

Al igual que en el caso escalar se demuestra que son cumplidas las desigualdades $R \geqslant \delta E$ y $R_1 R_2 \leqslant \frac{\Delta}{4} R$, donde

$$\sigma = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha}, \quad \Delta = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2}.$$

De la teoría general del método alternativo triangular (véase el § 1 del capítulo X) se deduce, que para el valor óptimo del parámetro $\omega = \omega_0 = 2/\sqrt{\delta\Delta}$ tienen lugar las desigualdades operacionales

$$\overset{\circ}{\gamma}_1 B \leqslant R \leqslant \overset{\circ}{\gamma}_2 B, \quad \overset{\circ}{\gamma}_1 > 0, \tag{11}$$

donde $\overset{\circ}{\gamma}_1 = \dfrac{\delta}{2(1+\sqrt{\eta})}$, $\overset{\circ}{\gamma}_2 = \dfrac{\delta}{4\sqrt{\eta}}$, $\eta = \dfrac{\delta}{\Delta}$. Comparando (7), (9) y (11), encontramos, que los operadores A y B satisfacen las desigualdades (9) con $\gamma_1 = c_1 \overset{\circ}{\gamma}_1$ y $\gamma_2 = c_2 \overset{\circ}{\gamma}_2$.

Utilizando para el esquema (8) el conjunto de los parámetros de Chebishev τ_k, obtendremos que el método iterativo alternativo triangular construído exige $n = O\left(\frac{1}{\sqrt{|h|}} \sqrt{\frac{c_2}{c_1}} \ln \frac{2}{\varepsilon}\right)$ iteraciones, donde $|h|^2 = h_1^2 + h_2^2 + \ldots + h_p^2$. Ya que el paso de y_k a y_{k+1} se realiza por fórmulas explícitas con un gasto de $O(m_0 N_1 N_2 \ldots N_p)$ operaciones aritméticas, el número total de operaciones para encontrar la solución del problema (4) con una exactitud ε, se estima por la cantidad

$$Q = O\left(m_0 N^{p+0,5} \sqrt{\frac{c_2}{c_1}} \ln \frac{1}{\varepsilon}\right),$$

si $l_1 = l_2 = \ldots = l_p$ y $N_1 = N_2 = \ldots = N_p = N$.

Como conclusión señalemos, que los métodos iterativos examinados más arriba convergen en el espacio energético H_D, donde en calidad del operador D se puede tomar uno de los operadores A, B ó $AB^{-1}A$.

2. Sistema de ecuaciones de la teoría de elasticidad. Examinemos el sistema de ecuaciones de la teoría estacionaria de elasticidad (ecuaciones de Lame)

$$Lu = \mu \Delta u + (\lambda + \mu) \operatorname{grad} \operatorname{div} u = -f(x), \tag{12}$$

donde $u = (u^1, u^2, \ldots, u^p)$, $f = (f^1, f^2, \ldots, f)$, $x = (x_1, x_2, \ldots, x_p)$, $\lambda > 0$ y $\mu > 0$ son las constantes de Lame.

Escribamos la ecuación (12) en la forma del sistema

$$(Lu)^s = \mu \sum_{\alpha=1}^{p} \frac{\partial^2 u^s}{\partial x_\alpha^2} + (\lambda+\mu) \sum_{\beta=1}^{p} \frac{\partial^2 u^\beta}{\partial x_\beta \partial x_s} = \\ = -f^s, \quad t=1, 2, \ldots, p. \qquad (13)$$

Para $p = 2$ el sistema (13) se puede escribir de modo siguiente

$$(\lambda+2\mu)\frac{\partial^2 u^1}{\partial x_1^2} + \mu \frac{\partial^2 u^1}{\partial x_2^2} + (\lambda+\mu)\frac{\partial^2 u^2}{\partial x_1 \partial x_2} = -f^1(x_1, x_2),$$
$$(\lambda+\mu)\frac{\partial^2 u^1}{\partial x_1 \partial x_2} + \mu \frac{\partial^2 u^2}{\partial x_1^2} + (\lambda+2\mu)\frac{\partial^2 u^2}{\partial x_2^2} = -f^2(x_1, x_2).$$

Este sistema describe el equilibrio de un cuerpo sólido elástico isotrópico y homogéneo en el caso de una deformación plana. Las funciones desconocidas $u^1(x_1, x_2)$ y $u^2(x_1, x_2)$ tienen el significado de las traslaciones de un punto por las direcciones de los ejes Ox_1 y Ox_2 respectivamente.

Para el sistema (12) puede ser planteado el problema de buscar un vector $u(x)$, que satisfaga la ecuación (12) en la región G y que tome valores dados sobre la frontera Γ

$$u(x) = g(x), \quad x \in \Gamma. \qquad (14)$$

Comparando (13) con (2), encontramos, que el sistema (12), (14) puede ser escrito en la forma (1), donde $m_0 = p$,

$$k_{\alpha\beta}^{sm} = \mu\sigma_{\alpha\beta}\sigma_{sm} + (\lambda+\mu)[\theta\sigma_{\alpha s}\sigma_{\beta m} + (1-\theta)\,\delta_{\alpha m}\delta_{\beta s}], \qquad (15)$$

y θ es una constante arbitraria, $\delta_{ij} = \begin{cases} 1, & i=j, \\ 0, & i \neq j. \end{cases}$ Realmente, sustituyendo (15) en (2), tendremos

$$(Lu)^s = \sum_{\alpha,\beta=1}^{p} \sum_{m=1}^{p} \frac{\partial}{\partial x_\alpha}\left(k_{\alpha\beta}^{sm}\frac{\partial u^m}{\partial x_\beta}\right) = \\ = \mu \sum_{\alpha,\beta=1}^{p} \sum_{m=1}^{p} \delta_{\alpha\beta}\delta_{sm}\frac{\partial^2 u^m}{\partial x_\alpha \partial x_\beta} + \\ + (\lambda+\mu)\left[\theta \sum_{\alpha,\beta=1}^{p} \sum_{m=1}^{p} \delta_{\alpha s}\delta_{\beta m}\frac{\partial^2 u^m}{\partial x_\alpha \partial x_\beta} + \right. \\ \left. + (1-\theta) \sum_{\alpha,\beta=1}^{p} \sum_{m=1}^{p} \delta_{\alpha m}\delta_{\beta s}\frac{\partial^2 u^m}{\partial x_\alpha \partial x_\beta}\right] =$$

$$= \mu \sum_{\alpha=1}^{p} \frac{\partial^2 u^s}{\partial x_\alpha^2} + (\lambda+\mu)\Big[\theta \sum_{\beta=1}^{p} \frac{\partial^2 u^\beta}{\partial x_s\, \partial x\lambda} +$$

$$+ (1-\theta) \sum_{\alpha=1}^{p} \frac{\partial^2 u^\alpha}{\partial x_\alpha\, \partial x_s}\Big] = \mu \sum_{\alpha=1}^{p} \frac{\partial^2 u^s}{\partial x_\alpha^2} + (\lambda+\mu) \sum_{\beta=1}^{p} \frac{\partial^2 u^\beta}{\partial x_s\, \partial x\beta}.$$

La afirmación está demostrada.

Hallemos ahora las constantes c_1 y c_2 en las desigualdades (3). Mostremos que $c_1 = \mu$. Tenemos

$$\sum_{s,\,m=1}^{p} \sum_{\alpha,\,\beta=1}^{p} k_\alpha^{sm}{}_\beta \xi_\alpha^s \xi_\beta^m = \mu \sum_{\alpha,\,s=1}^{p} (\xi_\alpha^s)^2 + (\lambda+\mu)\Big[\theta \sum_{\alpha,\,s=1}^{p} \xi_\alpha^\alpha \xi_s^s +$$

$$+ (1-\theta) \sum_{\alpha,\,s=1}^{p} \xi_\alpha^s \xi_s^\alpha\Big] = \mu \sum_{\alpha,\,s=1}^{p} (\xi_\alpha^s)^2 +$$

$$+ (\lambda+\mu)\,[\theta\,(\sum_{\alpha=1}^{p} \xi_\alpha^\alpha)^2 + (1-\theta) \sum_{\alpha,\,s=1}^{p} \xi_\alpha^s \xi_s^\alpha]. \qquad (16)$$

Poniendo aquí $\theta = 1$, hallaremos

$$\sum_{\alpha,\,\beta=1}^{p} (k_{\alpha\beta}\xi_\alpha,\ \xi_\beta) = \mu \sum_{\alpha=1}^{p} |\,\xi_\alpha\,|^2 +$$

$$+ (\lambda+\mu)\,(\sum_{\alpha=1}^{p} \xi_\alpha^\alpha)^2 \geqslant \mu \sum_{\alpha=1}^{p} |\,\xi_\alpha\,|^2.$$

No es difícil mostrar también, que $c_2 = \lambda + 2\mu$. Poniendo en (16) $\theta = 0$ y utilizando la desigualdad de Cauchy — Buniakovsky, obtendremos

$$\sum_{\alpha,\,\beta=1}^{p} (k_{\alpha\beta}\xi_\alpha,\ \xi_\beta) = \mu \sum_{\alpha=1}^{p} |\,\xi_\alpha\,|^2 +$$

$$+ (\lambda+\mu) \sum_{\alpha,\,s=1}^{p} \xi_\alpha^s \xi_s^\alpha \leqslant \mu \sum_{\alpha=1}^{p} |\,\xi_\alpha\,|^2 +$$

$$+ \frac{(\lambda+\mu)}{2}\Big[\sum_{\alpha,\,s=1}^{p} (\xi_\alpha^s)^2 + \sum_{\alpha,\,s=1}^{p} (\xi_s^\alpha)^2\Big] =$$

$$= \mu \sum_{\alpha=1}^{p} |\,\xi_\alpha\,|^2 + (\lambda+\mu) \sum_{\alpha,\,s=1}^{p} (\xi_\alpha^s)^2 = (\lambda+2\mu) \sum_{\alpha=1}^{p} |\,\xi_\alpha\,|^2.$$

Construyamos ahora un esquema de diferencias que aproxime el problema (12), (14). Sustituyendo (15) en el esquema de diferencias (4), tendremos

$$(\Lambda^- y)^s = \mu \sum_{\alpha=1}^{p} y^s_{\bar{x}_\alpha x_\alpha} + 0{,}5\,(\lambda+\mu) \sum_{\beta=1}^{p} (y^\beta_{\bar{x}_\beta x_s} + y^\beta_{x_\beta \bar{x}_s} = \\ = -\varphi^s, \quad x \in \omega. \qquad (17)$$

$$y^s(x) = g^s(x), \quad x \in \gamma, \quad s = 1, 2, \ldots, p,$$

donde $\bar{\omega} = \omega \cup \gamma$ es la red introducida en el punto 1.

Quedan por determinar los operadores A y R, tal como fue hecho en el punto 1. La condición de simetría (6) está cumplida, por eso, empleando la primera fórmula de Green de diferencias, encontramos que los operadores A y R son autoconjugados en H y tienen lugar las desigualdades $c_1 R \leqslant A \leqslant c_2 R$, donde $c_1 = \mu$ y $c_2 = \lambda + 2\mu$.

Aquí los razonamientos ulteriores coinciden con aquellos que se realizaron en el punto 1. Así, el método iterativo (8) con $B = R$ y los parámetros de Chebishev τ_k se caracteriza por la siguiente estimación para el número de iteraciones:

$$n \geqslant n_0(\varepsilon) = \frac{\ln 0{,}5\varepsilon}{\ln \rho_1}, \quad \rho_1 = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{c_1}{c_2} = \frac{\mu}{\lambda+2\mu},$$

y el método alternativo triangular, construído sobre la base de un regularizador R, se caracteriza por la misma estimación, donde

$$\xi = \frac{\mu}{\lambda+2\mu} \frac{2\sqrt{\eta}}{1+\sqrt{\eta}}, \quad \eta = \frac{\delta}{\Delta},$$

$$\delta = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2} \operatorname{sen}^2 \frac{\pi h_\alpha}{2l_\alpha}, \quad \Delta = \sum_{\alpha=1}^{p} \frac{4}{h_\alpha^2}.$$

De esta manera, para el método alternativo triangular el número de iteraciones es proporcional a $\sqrt{\frac{\lambda+2\mu}{\mu}} = \sqrt{2+\frac{\lambda}{\mu}}$:

$$n_0(\varepsilon) = \sqrt{2+\frac{\lambda}{\mu}}\, n_0^*(\varepsilon),$$

donde $n_0^*(\varepsilon)$ es el número de iteraciones para resolver la ecuación p-dimensional de Poisson en diferencias por el método alternativo triangular.

§ 4. Métodos de resolución de las ecuaciones elípticas en las regiones no regulares

1. Problemas de diferencias en las regiones de forma compuesta y métodos para resolverlos. En los capítulos anteriores fueron examinados los métodos fundamentales directos e iterativos a fin de resolver los problemas de diferencias de contorno para las ecuaciones elípticas. Detengámonos en la utilización de dichos métodos para resolver los problemas de diferencias en las regiones de forma compuesta.

Supongamos que en la región $\overline{G}$ hay que resolver un problema de contorno para la ecuación en diferencias elíptica. Son posibles varios enfoques para resolver el problema planteado. Si este problema se escribe en forma de un sistema de ecuaciones vectoriales tripuntuales, se puede emplear el método de factorización métrica, construido en el capítulo II. La inconveniencia principal de este método consiste en un gran volumen de información intermedia que debe memorizarse en el proceso de resolución. Al resolverse un problema único, la obtención de esta información ocupa la mayor parte del tiempo correspondiente a la resolución de todo el problema. Otra situación se crea si se resuelve una serie de problemas de diferencias que difieren únicamente por sus segundos miembros. Aquí la parte fundamental de los datos intermedios se halla sólo al resolver el primer problema y dicha información se utiliza para resolver cada problema posterior.

Aplicando los métodos iterativos, llegamos a la necesidad de resolver una serie de problemas. ¿Cómo elegir de mejor modo el operador B en un esquema iterativo, si el problema de diferencias viene dado en una región de forma compuesta? Es poco probable que se pueda dar alguna recomendación universal. En el caso cuando el operador de diferencias A es autoconjugado, se puede recomendar emplear el método alternativo triangular. En este método la conversión del operador B es un problema ordinario, pero el número de iteraciones depende de la cantidad de nodos en la red.

Recordemos que en el caso cuando $\overline{G}$ es un rectángulo y el operador A corresponde, por ejemplo, a un operador de diferencias $\Lambda y = \sum_{\alpha=1}^{2} (a_\alpha(x)\, y_{\bar{x}_\alpha})_{x_\alpha}$ con coeficientes variables, prefijado sobre la red uniforme, es racional en calidad de B

tomar un operador que corresponde al operador de Laplace de diferencia $\mathcal{A}y = \sum_{\alpha=1}^{2} y_{\bar{x}_\alpha x_\alpha}$. Entonces el número de iteraciones no depende del número de nodos en la red, y a fin de resolver la ecuación $By_{k+1} = F$ se pueden utilizar los métodos eficaces directos de reducción total o de separación de variables.

En el caso cuando la región $\bar{G}$ es distinta del rectángulo ¿qué da este enfoque? Aquí la principal dificultad consistirá en resolver la ecuación $By_{k+1} = F$. Como antes, el número de iteraciones no dependerá del número de nodos, por eso el volumen total de cálculos quedará definido, en lo fundamental, por el método aplicado para resolver la ecuación indicada. A pesar de que el operador B tiene una ordinaria estructura estándar, es imposible emplear inmediatamente los métodos directos para resolver la ecuación $By_{k+1} = F$. Es necesario efectuar un paso más, es decir, reducir este problema al problema prefijado sobre la red en algún rectángulo (o rectángulos).

En este párrafo examinaremos el método encaminado a resolver el problema de Dirichlet de diferencias para la ecuación de Poisson en la región de forma completa. En el primer caso la región inicial se cortará en dos rectángulos, en el otro, se completará hasta un rectángulo que acotará toda la región a examinar. El problema formulado se convertirá en uno o dos problemas que se resuelven sobre la red en el rectángulo. Dicho método de resolución del problema planteado permite utilizar los métodos directos de reducción total y de separación de variables para resolver el problema transformado.

2. **Método de separación de regiones.** Sea que en la región compuesta, constituida por varias regiones regulares, se necesita hallar la solución de un problema de contorno de diferencias. Vamos a suponer que en cada región regular la ecuación en diferencias dada, para ciertas condiciones de contorno determinadas, podrá resolverse sobre las fronteras de separación de estas regiones mediante uno de los métodos iterativos o directos efectivos. Partiendo de dichas suposiciones, para la región compuesta, el problema inicial se podrá resolver en dos etapas. Al principio, en las fronteras de separación de las regiones, se determinan las condiciones de contorno, las cuales satisfacerá la solución buscada, acto seguido en cada región regular se resuelve, independiente-

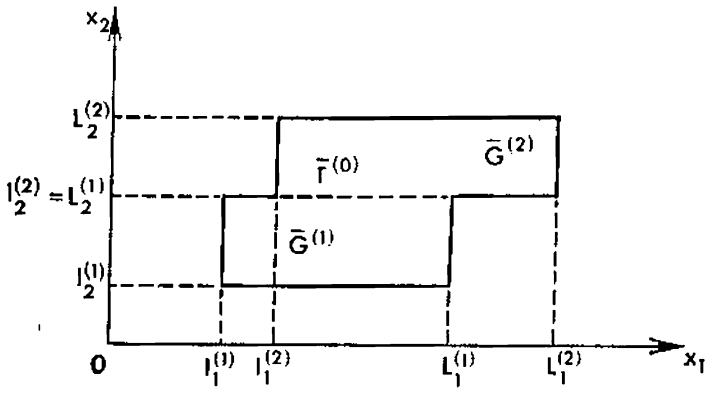

Fig. 4

mento, dicha ecuación en diferencias respecto a las correspondientes condiciones de contorno.

Expondremos este método recurriendo a un ejemplo en que el problema de Dirichlet de diferencias para la ecuación de Poisson se resuelve en una región compuesta de dos rectángulos colindantes por medio de un segmento recto perteneciente a los mismos. A fin de resolver el problema de diferencias se pueden utilizar los métodos directos construidos anteriormente en los capítulos III — IV.

Supongamos que la región $\overline{G}$ está compuesta de dos rectángulos $\overline{G} = \overline{G}^{(1)} \cup \overline{G}^{(2)}$: $\overline{G}^{(\alpha)} = \{l_1^{(\alpha)} \leqslant x_1 \leqslant L_1^{(\alpha)},\ l_2^{(\alpha)} \leqslant x_2 \leqslant L_2^{(\alpha)}\}$, $\alpha = 1, 2$, que tienen tangencia formando el segmento $\overline{\Gamma}^{(0)} = \{\max(l_1^{(1)}, l_1^{(2)}) \leqslant x_1 \leqslant \min(L_1^{(1)}, L_1^{(2)}),\ x_2 = l_2^{(2)} = L_2^{(1)}\}$ (véase la fig. 4).

Designemos con Γ la frontera de la región $\overline{G}$. Una parte de la frontera del rectángulo $\overline{G}^{(\alpha)}$ sin el segmento $\Gamma^{(0)}$ la designamos con $\Gamma^{(\alpha)}$, $\alpha = 1, 2$. Planteemos el siguiente problema: hallar en la región $\overline{G}$ la solución de la ecuación de Poisson que sobre Γ adquiera los valores prefijados

$$\begin{aligned} &Lu = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = -f(x), \quad x = (x_1, x_2) \in G, \\ &u(x) = g(x), \quad x \in \Gamma. \end{aligned} \tag{1}$$

Al principio vamos a formular el análogo de diferencias del problema (1). Para eso en el rectángulo $\overline{G}^{(\alpha)}$ introduci-

mos las red $\overline{\omega}^{(\alpha)}$ uniforme respecto de la dirección x_2:

$$\overline{\omega}^{(\alpha)} = \{x_{ij} = (x_1(i), \quad x_2(j)),\ x_1(i) = x_1(i-1) + \\ + h_1^{(\alpha)}(i), \quad 1 \leqslant i < N_1^{(\alpha)},$$

$$x_1(0) = l_1^{(\alpha)}, \quad x_1(N_1^{(\alpha)}) = L_1^{(\alpha)}, \quad x_2(j) = l_2^{(\alpha)} + jh_2^{(\alpha)},$$

$$0 \leqslant j \leqslant N_2^{(\alpha)}, \quad N_2^{(\alpha)} h_2^{(\alpha)} = L_2^{(\alpha)} - l_2^{(\alpha)}\}, \quad \alpha = 1,\ 2.$$

Las redes $\overline{\omega}^{(1)}$ y $\overline{\omega}^{(2)}$ deben ser concordadas. Esto significa que los nodos de dichas redes que se encuentran sobre $\overline{\Gamma}^{(0)}$ coinciden.

Para la región mostrada en la fig. 4 las redes concordadas pueden prefijarse de la siguiente manera. Sobre cada segmento

$$[l_1^{(1)},\ l_1^{(2)}], \quad [l_1^{(2)},\ L_1^{(1)}] \quad \text{y} \quad [L_1^{(1)},\ L_1^{(2)}]$$

se construye su propia red unidimensional, uniforme o no uniforme. A través de los nodos se trazan rectas paralelas al eje Ox_2. A continuación en los segmentos $[l_2^{(1)},\ L_2^{(1)}]$ y $[l_2^{(2)},\ L_2^{(2)}]$ se introducen las redes uniformes unidimensionales con pasos de $h_2^{(1)}$ y $h_2^{(2)}$, respectivamente, por los nodos de los cuales se trazan rectas paralelas al eje Ox_1. Los puntos de intersección de las rectas verticales y horizontales, que pertenecen al rectángulo $\overline{G}^{(\alpha)}$, forman la red $\overline{\omega}^{(\alpha)}$, $\alpha = 1,\ 2$.

La unión de las redes $\overline{\omega}^{(1)}$ y $\overline{\omega}^{(2)}$ forma la red $\overline{\omega}$ en la región $\overline{G}$. Designemos los nodos interiores de la red $\overline{\omega}$ con ω, y sus nodos de frontera con γ. Anotemos asimismo los nodos interiores de la red $\overline{\omega}^{(\alpha)}$ con $\omega^{(\alpha)}$, y sus nodos $\overline{\omega}^{(\alpha)}$ que se encuentran en $\Gamma^{(\alpha)}$ por $\gamma^{(\alpha)}$. Mediante $\gamma^{(0)}$ designemos los nodos comunes para las redes $\overline{\omega}^{(1)}$ y $\overline{\omega}^{(2)}$, que se encuentran en el segmento $\Gamma^{(0)}$. Entonces $\overline{\omega}^{(\alpha)} = \omega^{(\alpha)} \cup \gamma^{(\alpha)} \cup \gamma^{(0)}$, $\alpha = 1,\ 2$, y $\omega = \omega^{(1)} \cup \omega^{(2)} \cup \gamma^{(0)}$.

Sobre la red $\overline{\omega}$ construiremos un esquema de diferencias que aproxima el problema (1):

$$\Lambda y = -\varphi(x), \quad x \in \omega, \quad y(x) = g(x), \quad x \in \gamma, \tag{2}$$

donde $\Lambda = \Lambda_1 + \Lambda_2$, además los operadores Λ_1 y Λ_2 se determinan con ayuda de las fórmulas: $\Lambda_1 y = y_{\bar{x}_1 x_1}$ y

$$\Lambda_2 y = \begin{cases} y_{\bar{x}_2 x_2}, & x \in \omega^{(\alpha)}, \\ \dfrac{2}{h_2^{(1)} + h_2^{(2)}} \left[\dfrac{y(x_1, x_2 + h_2^{(2)}) - y(x_1, x_2)}{h_2^{(2)}} - \right. \\ \qquad \left. - \dfrac{y(x_1, x_2) - y(x_1, x_2 - h_2^{(1)})}{h_2^{(1)}} \right], & x \in \gamma^{(0)} \end{cases}$$

Pasemos a construir el método para resolver el problema (2). Señalemos que si se ha hallado el valor de $y(x)$ para $x \in \gamma^{(0)}$, el problema (2) se descompone en dos problemas de Dirichlet independientes para la ecuación de Poisson en los rectángulos $\overline{G}^{(1)}$ y $\overline{G}^{(2)}$. Por lo tanto, para el caso a examinar, en la frontera de separación de las regiones regulares (rectángulos) debe ser definida la condición de contorno del primer género.

La obtención de los valores buscados de $y(x)$ sobre $\gamma^{(0)}$ vamos a realizarla de la manera siguiente. Reduzcamos al principio el problema (2) al problema de condiciones de contorno homogéneas y cuyo segundo miembro difiere de cero únicamente cuando $x \in \gamma^{(0)}$. Para este fin determinemos la función $v(x)$ que es la solución del problema

$$\begin{aligned} &\Lambda v = -\varphi(x), \quad x \in \omega - \gamma^{(0)}, \quad v(x) = 0, \quad x \in \gamma^{(0)}, \\ &v(x) = g(x) \quad x \in \gamma. \end{aligned} \tag{3}$$

Subrayemos que el problema (3) se descompone en dos problemas dados sobre las redes $\overline{\omega}^{(1)}$ y $\overline{\omega}^{(2)}$. Hemos pasado del problema (2) al problema (3), sustituyendo la ecuación $\Lambda y = -\varphi(x)$ por la condición $v(x) = 0$ en los nodos pertenecientes a $\gamma^{(0)}$.

De (2), (3) encontramos que la función $z = y - v$ es la solución del problema

$$\begin{aligned} &\Lambda z = 0, \quad x \in \omega - \gamma^{(0)}, \quad \Lambda z = -\psi(x), \quad x \in \gamma^{(0)}, \\ &z(x) = 0, \quad x \in \gamma, \end{aligned} \tag{4}$$

donde $\varphi(x) = \varphi(x) + \Lambda v(x)$. Puesto que $v(x) = 0$ para $x \in \gamma^{(0)}$, entonces $y(x) = z(x)$ para $x \in \gamma^{(0)}$. Por consiguiente, la solución buscada del problema (2) podrá hallarse como la solución del problema de contorno

$$\begin{aligned} &\Lambda y = -\varphi(x), \quad x \in \omega - \gamma^{(0)}, \quad y(x) = z^{(0)}(x), \quad x \in \gamma^{(0)}, \\ &y(x) = g(x), \quad x \in \gamma, \end{aligned} \tag{5}$$

en que $z^{(0)}(x)$ no es más que el valor de la solución de $z(x)$ del problema (4) para $x \in \gamma^{(0)}$. El problema (5) lo mismo que el problema (3), se desintegra en dos problemas independientes en $\overline{G}^{(1)}$ y $\overline{G}^{(2)}$.

Encontremos $z^{(0)}(x)$. Para esto vamos a estudiar el problema (4) para todo $\psi(x)$. Puesto que Λ es el operador lineal de diferencias y la función $z(x)$ satisface la condición de contorno homogénea sobre γ, la solución del problema (4)

se expresa de manera lineal a través del segundo miembro de la ecuación (4). Tomando en consideración la estructura del segundo miembro, obtendremos que $z^{(0)}(x)$ y $\psi(x)$ están ligados por medio de la relación que se puede anotar en forma de

$$\sum_{\xi \in \gamma^{(0)}} a(x, \xi) z^{(0)}(\xi) = \psi(x), \quad x \in \gamma^{(0)}, \tag{6}$$

donde $a(x, \xi)$ son ciertos coeficientes.

Notemos que la regularidad de la matriz $\mathcal{A} = (a(x, \xi))$ del sistema (6) se deduce de la unicidad de solución del problema (4). Además, los coeficientes $a(x, \xi)$ no dependen de $\psi(x)$, y quedan definidos sólo mediante el operador Λ, la red $\bar{\omega}$ y la geometría de la región $\bar{G}$. Por lo tanto, el problema de encontrar $a(x, \xi)$ constituye una parte independiente del método que se examina. Si encontramos dichos coeficientes, entonces el valor buscado de $z^{(0)}(x)$ podrá calcularse resolviendo el sistema (6) para $\psi(x)$ indicado más arriba.

Examinemos el método de obtención de $a(x, \xi)$. De (4), (6) se desprende que para cualquier $\psi(x)$ la solución del problema (4) satisface las relaciones

$$\sum_{\xi \in \gamma^{(0)}} a(x, \xi) z^{(0)}(\xi) = -\Lambda z(x), \quad x \in \gamma^{(0)}. \tag{7}$$

Supongamos que $w(x)$ es la solución del problema

$$\begin{aligned} &\Lambda w = 0, \quad x \in \omega - \gamma^{(0)}, \quad w(x) = w^{(0)}(x), \quad x \in \gamma^{(0)}, \\ &w(x) = 0, \quad x \in \gamma \end{aligned} \tag{8}$$

para la función $w^{(0)}(x)$ arbitrariamente prefijada. A dicha solución le corresponde cierta función $\psi(x) = -\Lambda w(x)$, $x \in \gamma^{(0)}$. Por consiguiente, $w(x)$ es la solución del problema

$$\begin{aligned} &\Lambda w = 0, \quad x \in \omega - \gamma^{(0)}, \quad \Lambda w = -\psi(x), \quad x \in \gamma^{(0)}, \\ &w(0) = 0, \quad x \in \gamma. \end{aligned}$$

Debido a esto, en virtud de (7), son válidas las igualdades

$$\sum_{\xi \in \gamma^{(0)}} (a(x, \xi) w^{(0)}(\xi) = -\Lambda w(x), \quad x \in \gamma^{(0)}, \tag{9}$$

que permiten encontrar los coeficientes $a(x, \xi)$.

En efecto, pongamos en (8), (9)

$$w^{(0)}(x) = \delta(x - \eta) = \begin{cases} 0, & x \neq \eta, \\ 1, & x \neq \eta, \end{cases} \tag{10}$$

donde $x \in \gamma^{(0)}$ y η es un nodo arbitrario perteneciente a $\gamma^{(0)}$. De (9) obtendremos

$$a(x, \eta) = -\Lambda w(x), \quad x \in \gamma^{(0)}, \tag{11}$$

donde $w(x)$ es la solución del problema (8), (10). Eligiendo en calidad de η, sucesivamente, todos los nodos del conjunto $\gamma^{(0)}$, hallaremos los valores de $a(x, \xi)$ para todos x y ξ pertenecientes a $\gamma^{(0)}$.

De ese modo, para obtener la solución del problema (2) con ayuda del método construido más arriba, es necesario: 1) calcular los coeficientes $a(x, \xi)$ para $x, \xi \in \gamma^{(0)}$, 2) resolver el problema auxiliar (3) y calcular la función $\psi(x) = \varphi(x) + \Lambda v(x)$ y 3) resolver el sistema de ecuaciones algebraicas (6) y hallar la solución del problema (5).

OBSERVACIÓN. El método de separación de regiones puede aplicarse también en aquellos casos cuando sobre los lados $x_1 = l_1^{(\alpha)}$ y $x_1 = L_1^{(\alpha)}$, $\alpha = 1, 2$, así como sobre los lados $x_2 = l_2^{(1)}$ y $x_2 = L_2^{(2)}$ está prefijada cualquier combinación de las condiciones de contorno de primero, segundo o tercer género. Con este fin la condición de contorno de primer género, definida en los problemas (3) y (5) para $x \in \gamma$, debe ser sustituida por la correspondiente condición de contorno no homogénea sobre cada segmento de la frontera γ. En el problema (8) la condición de contorno homogénea de primer género sobre γ se cambia por la correspondiente condición homogénea.

En un segmento de la frontera Γ del rectángulo $\bar{G}$ que se encuentra en la recta $x_2 = l_2^{(2)} = L_2^{(1)}$ (véase la figura 4) podrá prefijarse una condición de contorno que no sea del primer género. Entonces en $\gamma^{(0)}$ se deben incluir aquellos nodos de la red $\bar{\omega}$ que pertenecen a dicho segmento de la frontera Γ.

3. **Algoritmo del método de separación de regiones.** Detengámonos a examinar con más detalles cada etapa del método trazado.

I. *Etapa preparatoria.* El objetivo de dicha etapa consiste en calcular $a(x, \xi)$. Si se resuelve una serie de problemas (2) con diferentes $\varphi(x)$ y $g(x)$, entonces será oportuno calcular los coeficientes $a(x, \xi)$ una vez y memorizarlos. El coeficiente $a(x, \eta)$ se halla mediante la fórmula (11), donde $w(x)$ es la solución del problema (8), (10) para $\eta \in \gamma^{(0)}$.

Al elegir el método para resolver dicho problema, conviene tomar en consideración que es suficiente hallar la función $w(x)$ sólo en los nodos situados sobre las líneas de

la red $\bar{\omega}$, más próximas por debajo y por encima de $\gamma^{(0)}$.

Designemos $w^{(\alpha)}(x) = w(x)$, $x \in \omega^{(\alpha)}$, $\alpha = 1, 2$. Utilizando estas designaciones, anotemos el problema (8), (10) en forma de dos problemas

$$\Lambda w^{(\alpha)} = 0, \quad x \in \omega^{(\alpha)}, \quad w^{(\alpha)}(x) = 0, \quad x \in \gamma^{(\alpha)},$$
$$w^{(\alpha)}(x) = \delta(x - \eta), \quad x, \ \eta \in \gamma^{(0)}, \quad \alpha = 1, 2. \tag{12}$$

Introducimos el vector $W_j^{(\alpha)}$, cuyas componentes son los valores $w^{(\alpha)}(x)$ en los nodos de la j-ésima línea de la red $\omega^{(\alpha)}$, y anotemos el problema (12) en forma del sistema de las ecuaciones vectoriales tripuntuales

$$-W_{j-1}^{(\alpha)} + C^{(\alpha)} W_j^{(\alpha)} - W_{j+1}^{(\alpha)} = 0, \quad 1 \leqslant j \leqslant N_2^{(\alpha)} - 1,$$
$$\alpha = 1, 2, \tag{13}$$

para el cual las condiciones de contorno se prefijan de la manera siguiente:

$$\begin{aligned} &W_0^{(1)} = 0, \quad W_{N_2^{(1)}}^{(1)} = F^{(1)}, \quad \text{si} \quad \alpha = 1, \\ &W_0^{(2)} = F^{(2)}, \quad W_{N_2^{(2)}}^{(2)} = 0, \quad \text{si} \quad \alpha = 2. \end{aligned} \tag{14}$$

Aquí $C^{(\alpha)}$ es una matriz de tres diagonales, correspondiente al operador de diferencias $2E + [h_2^{(\alpha)}]^2 \Lambda_1$, y los vectores $F^{(1)}$ y $F^{(2)}$ tienen una componente no nula.

Por la fórmula para la solución general de la ecuación vectorial tripuntual con coeficientes constantes que fue obtenida en el punto 3 del cap. I, hallemos los valores buscados para $W_{N_2^{(1)}-1}^{(1)}$ y $W_1^{(2)}$:

$$W_{N_2^{(1)}}^{(1)} = \left[U_{N_2^{(1)}-1}\left(\frac{C^{(1)}}{2}\right)\right]^{-1} U_{N_2^{(1)}-2}\left(\frac{C^{(1)}}{2}\right) F^{(1)},$$
$$W_1^{(2)} = \left[U_{N_2^{(2)}-1}\left(\frac{C^{(2)}}{2}\right)\right]^{-1} U_{N_2^{(2)}-2}\left(\frac{C^{(2)}}{2}\right) F^{(2)},$$

donde $U_n(t)$ es el polinomio de Chebishev de segundo género del grado n.

Empleando el lema 6 del cap. II para desarrollar la razón entre los polinomios $U_{n-2}(t)/U_{n-1}(t)$ en una suma

de fracciones elementales, obtenemos

$$W^{(1)}_{N_2^{(1)}-1} = \frac{2}{N_2^{(1)}} \sum_{h=1}^{N_2^{(1)}-1} \operatorname{sen}^2 \frac{k\pi}{N_2^{(1)}} \left(C^{(1)} - 2\cos\frac{k\pi}{N_2^{(1)}} E \right)^{-1} F^{(1)},$$

$$W^{(2)}_1 = \frac{2}{N_2^{(2)}} \sum_{h=1}^{N_2^{(2)}-1} \operatorname{sen}^2 \frac{k\pi}{N_2^{(2)}} \left(C^{(2)} - 2\cos\frac{k\pi}{N_1^{(2)}} E \right)^{-1} F^{(2)}.$$

Calculemos dichas sumas. Supongamos que $V_h^{(\alpha)}$ es la solución de la ecuación:

$$\left(C^{(\alpha)} - 2\cos\frac{k\pi}{N_2^{(\alpha)}} E \right) V_h^{(\alpha)} = \frac{2}{N_2^{(\alpha)}} \operatorname{sen}^2 \frac{k\pi}{N_2^{(\alpha)}} F^{(\alpha)},$$

$$1 \leqslant k \leqslant N_2^{(\alpha)} - 1, \qquad (15)$$

entonces

$$W^{(1)}_{N_2^{(1)}-1} = \sum_{h=1}^{N_2^{(1)}-1} V_h^{(1)}, \quad W^{(2)}_1 = \sum_{h=1}^{N_2^{(2)}-1} V_h^{(2)}. \qquad (16)$$

Pasemos en (15) y (16) de la anotación vectorial a la escalar. Obtendremos que sobre las respectivas líneas de la red $\bar{\omega}^{(\alpha)}$ las funciones $w^{(\alpha)}(x)$ se hallan mediante las fórmulas:

$$w^{(1)}(x_1, x_2) = \sum_{h=1}^{N_2^{(1)}-1} v_h^{(1)}(x_1), \quad l_1^{(1)} \leqslant x_1 \leqslant L_1^{(1)},$$

$$x_2 = L_2^{(1)} - h_2^{(1)}, \qquad (17)$$

$$w^{(2)}(x_1, x_2) = \sum_{h=1}^{N_2^{(2)}-1} v_h^{(2)}(x_1), \quad l_1^{(2)} \leqslant x_1 \leqslant L_1^{(2)},$$

$$x_2 = l_2^{(2)} + h_2^{(2)},$$

donde $v_h^{(\alpha)}(x_1)$ es la solución del problema tripuntual de contorno

$$\Lambda_1 v_h^{(\alpha)} - \lambda_h^{(\alpha)} v_h^{(\alpha)} = -\frac{2\delta(x-\eta)}{N_2^{(\alpha)} [h_2^{(\alpha)}]^2} \operatorname{sen}^2 \frac{k\pi}{N_2^{(\alpha)}},$$

$$l_1^{(\alpha)} < x_1 < L_1^{(\alpha)},$$

$$v_h^{(\alpha)}(l_1^{(\alpha)}) = v_h^{(\alpha)}(L_1^{(\alpha)}) = 0, \quad x, \eta \in \gamma^{(0)}, \qquad (18)$$

$$\lambda_h^{(\alpha)} = \frac{4}{[h_2^{(\alpha)}]} \operatorname{sen}^2 \frac{k\pi}{2N_2^{(\alpha)}}, \quad 1 \leqslant k \leqslant N_2^{(\alpha)} - 1.$$

Los coeficientes $a(x, \eta)$ se calculan según la fórmula (11) que para el caso examinado tiene la siguiente forma:

$$a(x, \eta) = \left(\frac{2}{h_2^{(1)} h_2^{(2)}} E - \Lambda_1\right) w^{(0)}(x) - \\ - \frac{2}{h_2^{(2)}[h_2^{(1)} + h_2^{(2)}]} w^{(2)}(x_1, x_2 + h_2^{(2)}) + \\ + \frac{2}{h_2^{(1)}[h_2^{(1)} + h_2^{(2)}]} w^{(1)}(x_1, x_2 - h_2^{(1)}),$$

$$x \in \gamma^{(0)}, \quad \eta \in \gamma^{(0)}, \tag{19}$$

donde $w^{(0)}(x) = \delta(x - \eta)$. Las fórmulas (17) — (19) describen por completo la etapa preparatoria.

Si el problema (18) se resuelve recurriendo al método de factorización, entonces para calcular todos los coeficientes $a(x, \xi)$ se necesitan $O(n(N_1^{(1)} N_2^{(1)} + N_1^{(2)} N_2^{(2)}))$ operaciones aritméticas, donde n es el número de los nodos que pertenecen a $\gamma^{(0)}$.

II. *Etapa auxiliar.* El objetivo de ésta consiste en obtener la solución del problema (3) en los nodos que se sitúan en las líneas de la red $\bar{\omega}$, más próximos por encima y por debajo de $\gamma^{(0)}$, así como calcular la función $\psi(x) = \varphi(x) + \Lambda v(x)$, donde $x \in \gamma^{(0)}$.

Designemos por medio de $v^{(\alpha)}(x)$ el valor de $v(x)$ para $x \in \bar{\omega}^{(\alpha)}$. De (3) obtendremos dos problemas

$$\Lambda v^{(\alpha)} = -\varphi(x), \quad x \in \omega^{(\alpha)}, \\ v^{(\alpha)}(x) = g(x), \quad x \in \gamma^{(\alpha)}, \quad v^{(\alpha)}(x) = 0, \\ x \in \gamma^{(0)}, \quad \alpha = 1, 2, \tag{20}$$

cada uno de los cuales puede ser resuelto por el método de reducción total que fue descrito en el cap. III. Supongamos que $N_2^{(1)}$ y $N_2^{(2)}$ son los grados de 2 y se emplea, por ejemplo, el primer algoritmo del método de reducción total (véase el punto 3 del § 2 del cap. III). Mostremos cómo se deben realizar los cálculos en el ejemplo del problema (20) para $\alpha = 2$. Al efectuar el uso directo del método de reducción se calculan y memorizan los vectores $p_j^{(k)}$. Al cumplir el uso inverso, sucesivamente, se calculan, pero no se memorizan, los vectores $V_j^{(2)}$ para $j = N_2^{(2)}/2, N_2^{(2)}/4, \ldots, 1$. Las componentes del último vector calculado dan los valores buscados de $v^{(2)}(x)$ para $l_1^{(2)} \leqslant x_1 \leqslant L_1^{(2)}$, $x_2 = l_2^{(2)} + h_2^{(2)}$.

De modo análogo se obtienen los valores de $v^{(1)}(x)$ para $l_1^{(1)} \leqslant x_1 \leqslant L_1^{(1)}$, $x_2 = L_2^{(1)} - h_2^{(1)}$. Utilizando los valores obtenidos, resulta fácil calcular la función $\psi(x)$ para $x \in \gamma^{(0)}$.

Para realizar esta etapa se requieren cumplir $O(N_1^{(1)} N_2^{(1)} \log_2 N_2^{(1)} + N_1^{(2)} N_2^{(2)} \log_2 N_2^{(2)})$ operaciones aritméticas.

III. *Etapa fundamental.* En esta etapa se halla la solución del problema inicial. Con este fin se resuelve el sistema (6) y se calcula $z^{(0)}(x)$. Después, empleando el método de reducción total se resuelve el problema (5) el cual puede ser anotado en forma de dos problemas

$$\begin{aligned} &\Lambda y^{(\alpha)} = -\varphi(x), \quad x \in \omega^{(\alpha)}, \\ &y^{(\alpha)}(x) = g(x), \quad x \in \gamma^{(\alpha)}, \quad y^{(\alpha)}(x) = z^{(0)}(x), \\ &x \in \gamma^{(0)}, \quad \alpha = 1, \ 2. \end{aligned} \tag{21}$$

La resolución de dichos problemas se empieza de inmediato por el uso inverso del método de reducción total, además se emplean los vectores memorizados $p_j^{(k)}$.

Si el sistema (6) se resuelve con ayuda del método de Gauss, entonces para eso se necesitarán $O(n^3)$ operaciones. Cabe señalar que si se resuelve una serie de los problemas (2), será racional ya en la etapa preparatoria cumplir la descomposición de la matriz $\mathcal{A}$ del sistema (6) en el producto $\mathcal{A} = LU$ de las matrices triangulares inferior y superior. En este caso, en la etapa fundamental la resolución del sistema (6) se cumplirá haciendo $O(n^2)$ operaciones. La solución de los problemas (21) se puede hallar mediante el método de reducción total, gastando $O(N_1^{(1)} N_2^{(2)} \log_2 N_2^{(1)} + N_1^{(2)} N_2^{(2)} \times \log_2 N_2^{(2)})$ operaciones aritméticas.

4. Método de completar la región hasta un rectángulo. Supongamos que sobre una región $\bar{G}$ acotada con la frontera Γ hay que resolver un problema de Dirichlet para la ecuación de Poisson (1).

Construyamos en la región $\bar{G}$ la red $\bar{\omega}$ Para este fin tracemos una familia de rectas $x_1 = x_1(i)$, $i = 0, \pm 1, \pm 2, \ldots$, y una familia de rectas equidistantes $x_2 = x_2(j)$, $j = 0, \pm 1, \pm 2, \ldots$ Los puntos $x_{ij} = (x_1(i), x_2(j))$ forman una red en el plano. Además, los puntos x_{ij} que pertenecen a G, forman un conjunto de los nodos interiores ω. Los puntos x_{ij} que pertenecen a Γ, así como los puntos de intersección de Γ con cualquiera recta de las familias trazadas, forman

el conjunto de los nodos de frontera γ. La red $\bar{\omega}$ es la unión de ω y γ, $\bar{\omega} = \omega \cup \gamma$.

Los nodos interiores de la red $\bar{\omega}$ los dividiremos en regulares y no regulares. El nodo $x_{ij} \in \omega$ se llama regular si los cuatro puntos de la red $x_{i\pm1 j}$, $x_{ij\pm1}$ adyacentes al primero pertenecen a ω, en caso contrario, será no regular. El conjunto de los nodos regulares lo designaremos con $\overset{\circ}{\omega}$, mientras que el de los no regulares, con $\overset{*}{\omega}$.

El esquema de diferencias que aproxima (1) se construye de modo ordinario. El operador diferencial L se sustituye por un operador de diferencias de cinco puntos Λ en cada nodo ω, mientras que en los nodos de frontera γ se prefija la condición de contorno de primer género. Notemos que el molde del operador Λ en el nodo regular está constituido por los nodos de la red, mientras que en el nodo no regular dicho molde contiene aunque sea un nodo que pertenece a γ, sin embargo, éste no es obligatoriamente un punto de la red introducida.

Nuestra tarea consiste en anotar las ecuaciones en diferencias de tal modo que relacionen entre sí las incógnitas sólo en los nodos de la red. Para eso es menester transformar las ecuaciones en diferencias en los nodos no regulares, trasladando al segundo miembro los valores de la solución buscada, conocidos sobre γ. En este caso variará la estructura del operador de diferencias Λ en cada nodo no regular. El operador Λ transformado en los puntos $\overset{*}{\omega}$ lo designaremos con $\overset{*}{\Lambda}$. Debido a esto resultará construido el esquema de diferencias

$$\begin{aligned} \Lambda y &= -\varphi(x), \quad x \in \overset{\circ}{\omega}, \\ \overset{*}{\Lambda} y &= -\overset{*}{\varphi}(x), \quad x \in \overset{*}{\omega}, \end{aligned} \tag{22}$$

donde $\Lambda y = y_{\overline{x}_1 \hat{x}_1} + y_{\overline{x}_2 x_2}$.

Pasemos a la construcción del método para resolver el problema (22). Sea $\bar{G}_0$ un rectángulo que contiene $\bar{G}$, los lados del cual están formados por las rectas de las familias de rectas construidas más arriba. Con $\bar{\omega}_0$ designemos el conjunto de los puntos x_{ij} de la red introducida anteriormente que pertenecen a $\bar{G}_0$. Es evidente que $\omega = \overset{\circ}{\omega} \cup \overset{*}{\omega}$ es parte de la red $\bar{\omega}_0$. Con γ_0 designaremos la frontera de la red $\bar{\omega}_0$.

De este modo, ω_0 es una red rectangular uniforme respecto a la dirección x_2, sobre el rectángulo $\overline{G}_0$, en una parte de nodos de la cual están dadas las ecuaciones en diferencias (22).

Observemos que el problema $\Lambda u = f(x)$, $x \in \omega_0$, $u(x) = 0$ para $x \in \gamma_0$, donde Λ es el operador de diferencias anteriormente definido, podrá ser resuelto con eficacia utilizando uno de los métodos directos trazados en los capítulos III y IV. Nuestro objetivo consiste en reducir el esquema (22) al problema indicado sobre la red ω_0 con el segundo miembro $f(x)$ especialmente elegido.

Planteemos el problema: elegir la función $\psi(x)$ de modo que la solución del problema

$$\begin{aligned} &\Lambda y = -\varphi(x), \quad x \in \overset{\circ}{\omega}, \quad \Lambda y = -\psi(x), \quad x \in \overset{*}{\omega}, \\ &\Lambda y = 0, \quad x \in \omega_0 - \omega, \quad y(x) = 0, \quad x \in \gamma_0 \end{aligned} \tag{23}$$

satisfaga la igualdad

$$\overset{*}{\Lambda} y = -\overset{*}{\varphi}(x), \quad x \in \overset{*}{\omega}. \tag{24}$$

Si se halla la función $\psi(x)$, entonces las soluciones de los problemas (22) y (23) coinciden en los nodos ω.

Para buscar la función $\psi(x)$ procederemos del modo siguiente. Primeramente reducimos (23) al problema, cuyo segundo miembro difiera del cero únicamente sobre $\overset{*}{\omega}$. Con este fin, definimos la función auxiliar $v(x)$ que es la solución del problema

$$\begin{aligned} &\Lambda v = -\varphi(x), \quad x \in \overset{\circ}{\omega}, \quad \Lambda v = 0, \quad x \in \omega_0 - \overset{\circ}{\omega}, \\ &v(x) = 0, \quad x \in \gamma_0. \end{aligned} \tag{25}$$

Notemos que la función $z = y - v$ es la solución del problema que satisface las exigencias formuladas

$$\begin{aligned} &\Lambda z = 0, \quad x \in \overset{\circ}{\omega}, \quad \Lambda z = -\psi(x), \quad x \in \overset{*}{\omega}, \\ &\Delta z = 0, \quad x \in \omega_0 - \omega, \quad z(x) = 0, \quad x \in \gamma_0. \end{aligned} \tag{26}$$

Además, si $\psi(x)$ es la función buscada, entonces de (23) — (25) obtendremos

$$\psi^*(x) = -\Lambda^* z = \varphi^*(x) + \Lambda^* v(x), \quad x \in \overset{*}{\omega}. \tag{27}$$

Señalemos que en virtud de la linealidad del operador Λ, homogeneidad de la condición de contorno y de la estructura del segundo miembro del problema (26), la solución $z(x)$ se

expresa linealmente por medio de $\psi(x)$. Ya que Λ^* es también un operador lineal, entonces de (27) se desprende que $\psi(x)$ y $\psi^*(x)$ deberán estar relacionadas mediante una razón lineal. Puesto que las funciones $\psi(x)$ y $\psi^*(x)$ están prefijadas en un mismo conjunto $\overset{*}{\omega}$, dicha razón tiene la forma

$$\sum_{\xi \in \overset{*}{\omega}} a(x, \xi)\, \psi(\xi) = \psi^*(x), \quad x \in \overset{*}{\omega}, \tag{28}$$

donde $a(x, \xi)$ son ciertos coeficientes. Si encontramos dichos coeficientes, entonces la solución del problema inicial (22) consistirá en obtener la solución del problema auxiliar (25) en calcular la función $\psi^*(x)$ según la fórmula (27), en resolver el sistema de las ecuaciones algebraicas (28) para determinar $\psi(x)$ y en resolver el problema (23) para encontrar $y(x)$ en los nodos de la red ω.

Los coeficientes $a(x, \xi)$ se hallan con ayuda de un procedimiento análogo al expuesto en el punto 2. Supongamos que $w(x)$ es la solución del problema

$$\Lambda w = 0, \quad x \in \overset{\circ}{\omega}, \quad \Lambda w = -\delta(x - \eta), \quad x \in \overset{*}{\omega},$$
$$\Lambda w = 0, \quad x \in \omega_0 - \omega, \quad w(x) = 0, \quad x \in \gamma_0,$$

donde η es un nodo perteneciente a $\overset{*}{\omega}$. Entonces

$$a(x, \eta) = -\Lambda^* w(x), \quad x \in \overset{*}{\omega}. \tag{29}$$

Eligiendo en calidad de η sucesivamente todos los puntos $\overset{*}{\omega}$, obtendremos según la fórmula (29) todos los coeficientes $a(x, \xi)$. Con esto finalizamos la construcción del método que permite resolver el problema de Dirichlet de diferencias para la ecuación de Poisson sobre una región arbitraria.

Capítulo XV

Métodos de resolución de ecuaciones elípticas en coordenadas curvilíneas ortogonales

En este capítulo se examinan ejemplos de resolución de problemas de diferencias, los cuales aproximan problemas de contorno para ecuaciones elípticas en sistemas de coordenadas curvilíneas. Para los problemas en sistema de coordenadas cilíndricas y polares se aclaran las condiciones de aplicabilidad de los métodos directos e iterativos, en particular, del método de las direcciones variables.

En el § 1 se cita el planteamiento de los problemas de contorno para las ecuaciones diferenciales. El párrafo 2 está dedicado a la exposición de los métodos directos e iterativos de resolución de problemas de diferencias en la (r, z)—geometría, y también de los problemas sobre la superficie de un cilindro. En el § 3 son examinados los métodos de resolución de problemas de diferencias en un círculo, un anillo y en un sector anular.

§ 1. Planteamiento de los problemas de contorno para las ecuaciones diferenciales

1. Ecuaciones elípticas en un sistema cilíndrico de coordenadas. Sea dada la ecuación de Poisson

$$Lu = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \frac{\partial^2 u}{\partial x_3^2} = -f(x), \quad x = (x_1, x_2, x_3). \quad (1)$$

Si para esta ecuación se plantea el problema de buscar la solución en un cilindro circular finito o en un tubo circular, entonces es natural examinarlo en coordenadas cilíndricas. En este sistema de coordenadas la ecuación de Poisson (1) tiene la forma

$$L_{r\varphi z}u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \varphi^2} + \frac{\partial^2 u}{\partial z^2} =$$
$$= -f(r, \varphi, z), \quad (2)$$

donde $r = \sqrt{x_1^2 + x_2^2}$, $\operatorname{tg}\varphi = x_2/x_1$, $z = x_3$.

La ecuación (1) describe, por ejemplo, la distribución estacionaria de temperatura $u = u(x_1, x_2, x_3)$ en un medio homogéneo. Si el medio no es homogéneo, pero es isotrópico, entonces en lugar de (1) se debe examinar la ecuación

$$Lu = \operatorname{div}(k \operatorname{grad} u) = \sum_{\alpha=1}^{3} \frac{\partial}{\partial x_\alpha}\left(k(x)\frac{\partial u}{\partial x_\alpha}\right) = -f(x), \qquad (3)$$

a la cual le corresponde en el (r, φ, z) sistema la ecuación

$$L_{r\varphi z}u = \frac{1}{r}\frac{\partial}{\partial r}\left(rk\frac{\partial u}{\partial r}\right) + \frac{1}{r}\frac{\partial}{\partial \varphi}\left(k\frac{\partial u}{\partial \varphi}\right) + \frac{\partial}{\partial z}\left(k\frac{\partial u}{\partial z}\right) = -f. \qquad (4)$$

Si el medio es anisotrópico, es decir, si el coeficiente de conductibilidad térmica depende no solamente del punto, sino también de la dirección, entonces en lugar de (3) tendremos la ecuación con derivadas mixtas

$$Lu = \sum_{\alpha,\,\beta=1}^{3}{}' \frac{\partial}{\partial x_\alpha}\left(k_{\alpha\beta}\frac{\partial u}{\partial x_\beta}\right) = -f(x). \qquad (5)$$

A la ecuación (5) en un sistema cilíndrico de coordenadas le corresponde la ecuación

$$\begin{aligned} L_{r\varphi z}u = {} & \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(\bar{k}_{11}\frac{\partial u}{\partial r} + \frac{\bar{k}_{12}}{r}\frac{\partial u}{\partial \varphi} + \bar{k}_{13}\frac{\partial u}{\partial z}\right)\right] + \\ & + \frac{1}{r}\frac{\partial}{\partial \varphi}\left(\bar{k}_{21}\frac{\partial u}{\partial r} + \frac{\bar{k}_{22}}{r}\frac{\partial u}{\partial \varphi} + \bar{k}_{23}\frac{\partial u}{\partial z}\right) + \\ & + \frac{\partial}{\partial z}\left(\bar{k}_{31}\frac{\partial u}{\partial r} + \frac{\bar{k}_{32}}{r}\frac{\partial u}{\partial \varphi} + \bar{k}_{33}\frac{\partial u}{\partial z}\right) = -f(r, \varphi, z), \qquad (6) \end{aligned}$$

donde los coeficientes $\bar{k}_{\alpha\beta}$ se expresan mediante $k_{\alpha\beta}$ por las fórmulas:

$$\bar{k}_{11} = k_{11}\cos^2\varphi + (k_{12} + k_{21})\operatorname{sen}\varphi\cos\varphi + k_{22}\operatorname{sen}^2\varphi,$$

$$\bar{k}_{12} = k_{12}\cos^2\varphi + (k_{22} - k_{11})\operatorname{sen}\varphi\cos\varphi - k_{21}\operatorname{sen}^2\varphi,$$

$$\bar{k}_{21} = k_{21}\cos^2\varphi + (k_{22} - k_{11})\operatorname{sen}\varphi\cos\varphi - k_{12}\operatorname{sen}^2\varphi,$$

$$\bar{k}_{22} = k_{11}\operatorname{sen}^2\varphi - (k_{12} + k_{21})\operatorname{sen}\varphi\cos\varphi + k_{22}\cos^2\varphi,$$

$$\bar{k}_{13} = k_{13}\cos\varphi + k_{23}\operatorname{sen}\varphi, \quad \bar{k}_{23} = k_{23}\cos\varphi - k_{13}\operatorname{sen}\varphi,$$

$$\bar{k}_{31} = k_{31}\cos\varphi + k_{32}\operatorname{sen}\varphi, \quad \bar{k}_{32} = k_{32}\cos\varphi - k_{31}\operatorname{sen}\varphi, \quad \bar{k}_{33} = k_{33}$$

La ecuación (6) se llama *ecuación con derivadas mixtas en un sistema cilíndrico de coordenadas*. Si $\overline{k}_{\alpha\beta} = 0$ para $\alpha \neq \beta$, entonces (6) toma la forma

$$L_{r\varphi z}u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\overline{k}_1\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \varphi}\left(\overline{k}_2\frac{\partial u}{\partial \varphi}\right) + \\ + \frac{\partial}{\partial z}\left(\overline{k}_3\frac{\partial u}{\partial z}\right) = -f, \quad (7)$$

donde $\overline{k}_\alpha = \overline{k}_{\alpha\alpha}$, $\alpha = 1, 2, 3$, y se llama ecuación sin derivadas mixtas.

Señalemos, que si $k_{\alpha\beta} = k_{\beta\alpha}$, entonces $\overline{k}_{\alpha\beta} = \overline{k}_{\beta\alpha}$ y viceversa. Las ecuaciones (2) y (4) citadas más arriba son casos particulares de la ecuación (7), correspondientes a $\overline{k}_\alpha \equiv 1$ y $\overline{k}_\alpha \equiv k$.

La ecuación (5) será fuertemente elíptica, si existe una constante $c_1 > 0$ tal, que para cualesquiera ξ_1, ξ_2 y ξ_3 es válida la desigualdad

$$\sum_{\alpha,\beta=1}^{3} k_{\alpha\beta}(x)\,\xi_\alpha\xi_\beta \geqslant c_1 \sum_{\alpha=1}^{3}\xi_\alpha^2. \quad (8)$$

Si en (8) hacemos un cambio, poniendo

$$\xi_1 = \overline{\xi}_1\cos\varphi - \overline{\xi}_2\operatorname{sen}\varphi, \quad \xi_2 = \overline{\xi}_1\operatorname{sen}\varphi + \overline{\xi}_2\cos\varphi, \quad \xi_3 = \overline{\xi}_3,$$

entonces la desigualdad (8) toma la forma

$$\sum_{\alpha,\beta=1}^{3} \overline{k}_{\alpha\beta}\overline{\xi}_\alpha\overline{\xi}_\beta \geqslant c_1 \sum_{\alpha=1}^{3}\xi_\alpha^2. \quad (9)$$

En la práctica se encuentran con más frecuencia dos casos.

A) En el caso de simetría axial los coeficientes y el segundo miembro de la ecuación y también la misma solución no dependen del ángulo φ. Con esto, la ecuación (6) se simplifica:

$$L_{rz}u = \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(\overline{k}_{11}\frac{\partial u}{\partial r} + \overline{k}_{13}\frac{\partial u}{\partial z}\right)\right] + \\ + \frac{\partial}{\partial z}\left(\overline{k}_{31}\frac{\partial u}{\partial r} + \overline{k}_{33}\frac{\partial u}{\partial r}\right) = -f(r, z), \quad (10)$$

y en ausencia de derivadas mixtas la ecuación, correspondiente a (7), tiene la forma

$$L_{rz}u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\overline{k}_1\frac{\partial u}{\partial r}\right) + \frac{\partial}{\partial z}\left(\overline{k}_3\frac{\partial u}{\partial z}\right) = -f(r, z). \quad (11)$$

B) En el caso plano los coeficientes, el segundo miembro y la solución de la ecuación (6) no dependen de z y, por consiguiente, la ecuación (6) toma la forma

$$L_{r\varphi}u = \frac{1}{r}\frac{\partial}{\partial r}\left[r\left(\bar{k}_{11}\frac{\partial u}{\partial r} + \frac{\bar{k}_{12}}{r}\frac{\partial u}{\partial \varphi}\right)\right] + \\ + \frac{1}{r}\frac{\partial}{\partial \varphi}\left(\bar{k}_{21}\frac{\partial u}{\partial r} + \frac{\bar{k}_{22}}{r}\frac{\partial u}{\partial \varphi}\right) = -f(r, \varphi). \quad (12)$$

Si las derivadas mixtas están ausentes, entonces la ecuación tiene la forma

$$L_{r\varphi}u = \frac{1}{r}\frac{\partial}{\partial r}\left(r\bar{k}_1\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \varphi}\left(\bar{k}_2\frac{\partial u}{\partial \varphi}\right) = -f(r, \varphi). \quad (13)$$

En el caso plano se dice, que (12) y (13) son ecuaciones elípticas en el sistema polar de coordenadas.

Señalemos que cuando $\bar{k}_\alpha = 1$, $\alpha = 1, 2, 3$ las fórmulas (11) y (13) describen la ecuación de Poisson en los (r, z) y (r, φ) sistemas de coordenadas.

Algunas veces se exige resolver la ecuación de Poisson o una ecuación elíptica más general sobre la superficie de un cilindro de radio R. En este caso

$$L_{\varphi z} = \frac{1}{R}\frac{\partial}{\partial \varphi}\left(\frac{\bar{k}_{22}}{R}\frac{\partial u}{\partial \varphi} + \bar{k}_{23}\frac{\partial u}{\partial z}\right) + \\ + \frac{\partial}{\partial z}\left(\frac{\bar{k}_{32}}{R}\frac{\partial u}{\partial \varphi} + \bar{k}_{33}\frac{\partial u}{\partial z}\right) = -f(\varphi, z), \quad (14)$$

y la ecuación (7) sin derivadas mixtas toma la forma

$$L_{\varphi z}u = \frac{1}{R^2}\frac{\partial}{\partial \varphi}\left(\bar{k}_2\frac{\partial u}{\partial \varphi}\right) + \frac{\partial}{\partial z}\left(\bar{k}_3\frac{\partial u}{\partial z}\right) = -f(\varphi, z). \quad (14')$$

Señalemos, que el cambio $\varphi' = R\varphi$ permite reducir estas ecuaciones a ecuaciones elípticas usuales con coeficientes variables.

2. Problemas de contorno para ecuaciones en el sistema cilíndrico de coordenadas. Examinemos primeramente el *caso de simetría axial*. Ya que la solución no depende del ángulo φ, entonces en coordenadas cilíndricas la (r, z) región, donde se busca la solución, es el rectángulo $\overline{G} = \{l_1 \leqslant r \leqslant L_1,\ l_3 \leqslant z \leqslant L_3,\ l_1 \geqslant 0\}$. Si la región inicial representa un cilindro circular (hueco), entonces $l_1 > 0$.

Planteemos los problemas de contorno para la ecuación (10) en un rectángulo $\overline{G}$. En la región G se da la ecuación (10), sobre los lados $r = L_1$, $z = l_3$ y $z = L_3$, se prefija

una de las condiciones de contorno de primero, segundo o tercer género. Por ejemplo, las condiciones de contorno de tercer género tienen la forma:

$$\begin{aligned} -\overline{k}_{11}\frac{\partial u}{\partial r}-\overline{k}_{13}\frac{\partial u}{\partial z}&=\varkappa_1^+u-g_1^+(z), \quad r=L_1,\\ \overline{k}_{31}\frac{\partial u}{\partial r}+\overline{k}_{33}\frac{\partial u}{\partial z}&=\varkappa_3^-u-g_3^-(r), \quad z=l_3,\\ -\overline{k}_{31}\frac{\partial u}{\partial r}-\overline{k}_{33}\frac{\partial u}{\partial z}&=\varkappa_3^+u-g_3^+(r), \quad z=L_3. \end{aligned} \tag{15}$$

Para $l_1 = 0$ la ecuación (10) tiene una singularidad sobre el eje $r = 0$. En este caso se interesan, de ordinario, por las soluciones acotadas. Si $l_1 > 0$, entonces sobre el lado $r = l_1$ puede estar dada una de las condiciones de contorno de primero, segundo o tercer género. Por ejemplo, la condición de contorno de tercer género tiene la forma

$$\overline{k}_{11}\frac{\partial u}{\partial r}+\overline{k}_{13}\frac{\partial u}{\partial z}=\varkappa_1^-u-g_1^-(z), \quad r=l_1>0. \tag{16}$$

Si $l_1 = 0$, entonces una solución acotada se separa por la condición

$$\lim_{r\to 0} r\left(\overline{k}_{11}\frac{\partial u}{\partial r}+\overline{k}_{13}\frac{\partial u}{\partial z}\right)=0. \tag{17}$$

En las condiciones (15) y (16) $\varkappa_i^{\pm}(z)$ y $\varkappa_3^{\pm}(r)$ son las funciones no negativas. Si sobre la frontera del rectángulo $\overline{G}$ están prefijadas las condiciones de contorno de segundo género ($\varkappa_\alpha^{\pm} \equiv 0$), entonces el problema (10), (15) y (16) es soluble solamente al cumplirse la condición

$$\int_{l_1}^{L_1}\int_{l_3}^{L_3} rf(r, z)\,dr\,dz+\int_{l_3}^{L_3}[L_1g_1^+(z)+l_1g_1^-(z)]\,dz+ \\ +\int_{l_1}^{L_1} r\,[g_3^+(r)+g_3^-(r)]\,dr=0. \tag{18}$$

En este caso la solución no es única y se determina con exactitud hasta una constante, es decir, $u(r, z) = u_0(r, z) + \text{const}$, donde $u_0(r, z)$ es alguna solución.

Examinemos ahora la *ecuación* (14) *sobre la superficie de un cilindro*. En las coordenadas (φ, z) la región, donde se busca la solución, es el rectángulo $\overline{G} = \{l_2 \leqslant \varphi \leqslant L_2, l_3 \leqslant z \leqslant L_3, L_2 - l_2 \leqslant 2\pi\}$.

Sobre los lados $z = l_3$ y $z = L_3$ pueden ser dadas condiciones de contorno de primero, segundo o tercer género, por ejemplo

$$\begin{aligned} &\frac{1}{R}\overline{k}_{32}\frac{\partial u}{\partial \varphi}+\overline{k}_{33}\frac{\partial u}{\partial z}=\varkappa_3^- u-g_3^-(\varphi), \quad z=l_3,\\ &-\frac{1}{R}\overline{k}_{32}\frac{\partial u}{\partial \varphi}-\overline{k}_{33}\frac{\partial u}{\partial z}=\varkappa_3^+ u-g_3^+(\varphi), \quad z=L_3. \end{aligned} \tag{19}$$

Las condiciones de contorno de tal género pueden ser prefijadas sobre los lados $\varphi = l_2$ y $\varphi = L_2$, si la superficie no es cerrada ($L_2 - l_2 < 2\pi$). Por ejemplo, las condiciones de contorno de tercer género tienen la forma

$$\begin{aligned} &\frac{1}{R^2}\overline{k}_{22}\frac{\partial u}{\partial \varphi}+\frac{1}{R}\overline{k}_{23}\frac{\partial u}{\partial z}=\varkappa_2^- u-g_2^-(z), \quad \varphi=l_2,\\ &-\frac{1}{R^2}\overline{k}_{22}\frac{\partial u}{\partial \varphi}-\frac{1}{R}\overline{k}_{23}\frac{\partial u}{\partial z}=\varkappa_2^+ u-g_2^+(z), \quad \varphi=L_2. \end{aligned} \tag{20}$$

Aquí $\varkappa_2^{\pm}(z) \geqslant 0$ y $\varkappa_3^{\pm}(\varphi) \geqslant 0$.

La condición de solubilidad del problema (14), (19), (20) para $\varkappa_\alpha^{\pm} \equiv 0$ tiene la forma

$$\int_{l_2}^{L_2}\int_{l_3}^{L_3} f(\varphi, z)\,d\varphi\,dz+\int_{l_3}^{L_3}[g_2^+(z)+g_2^-(z)]\,dz+$$

$$+\int_{l_2}^{L_2}[g_3^+(\varphi)+g_3^-(\varphi)\,d\varphi=0.$$

Si la superficie es cerrada ($L_2 - l_2 = 2\pi$), entonces los lados $\varphi = l_2$ y $\varphi = L_2$ se identifican y se plantea el problema de hallar una solución periódica con período 2π de la ecuación (14), que satisfaga sobre los lados $z = l_3$ y $z = L_3$ una de las condiciones enumeradas más arriba. Si en dicho caso para las condiciones (19) $\varkappa_3^{\pm} \equiv 0$, entonces la condición de solubilidad (con exactitud hasta una constante) del problema indicado tiene la forma

$$\int_{l_2}^{L_2}\int_{l_3}^{L_3} f(\varphi, z)\,d\varphi\,dz+\int_{l_3}^{L_3}[g_3^+(\varphi)+g_3^-(\varphi)]\,d\varphi=0.$$

Formulemos ahora *los planteamientos de los problemas de contorno para la ecuación (12), definida en un sistema de coordenadas polares*, en el caso, cuando la región que se examina en variables cartesianas es un círculo, un anillo

o un sector anular. A las regiones indicadas en las coordenadas (r, φ) les corresponde el rectángulo $\overline{G} = \{l_1 \leqslant r \leqslant L_1,\ l_2 \leqslant \varphi \leqslant L_2,\quad l_1 \geqslant 0,\quad L_2 - l_2 \leqslant 2\pi\}$.

Supongamos primeramente que *la región inicial es un círculo*. La ecuación (12) se da en G; para $r = L_1$ se pone una de las condiciones de contorno de primero, segundo o tercer género. Por ejemplo, la condición de contorno de tercer género tiene la forma

$$-\overline{k}_{11}\frac{\partial u}{\partial r} - \frac{\overline{k}_{12}}{r}\frac{\partial u}{\partial \varphi} = \varkappa_1^+ u - g_1^+(\varphi), \quad r = L_1. \tag{21}$$

Para la corrección del problema (12), (21) es necesario imponer una condición complementaria en el centro del círculo. Por lo común se busca una solución acotada para $r = 0$. Esta solución satisface la condición

$$\lim_{r \to 0} r\left(\overline{k}_{11}\frac{\partial u}{\partial r} + \frac{\overline{k}_{12}}{r}\frac{\partial u}{\partial \varphi}\right) = 0. \tag{22}$$

En vista de que en un sistema de coordenadas polares el punto $r = 0$ del plano (x_1, x_2) tiene una coordenada φ arbitraria, entonces todos los puntos del lado del rectángulo $\overline{G}$ para $r = 0$ se identifican. Con esto $u(0, \varphi) = u_0 = \text{const}$ para $l_2 \leqslant \varphi \leqslant L_2$ en virtud de la continuidad de la solución.

A continuación, los lados $\varphi = l_2$ y $\varphi = L_2$ se identifican y se plantea el problema de hallar la solución periódica con período 2π de la ecuación (12), que satisface las condiciones enumeradas más arriba.

En el caso, cuando para $r = L_1$ está dada la condición de contorno de segundo género (21) con $\varkappa_1^+(\varphi) \equiv 0$, entonces la solución del problema existe, si se cumple la condición

$$\int_0^{2\pi}\int_0^{L_1} rf(r, \varphi)\, dr\, d\varphi + L_1 \int_0^{2\pi} g_1^+(\varphi)\, d\varphi = 0. \tag{23}$$

Al mismo tiempo la solución no es única y está definida con exactitud hasta una constante.

Sea ahora *la región inicial un anillo*, es decir, $l_1 > 0$. En este caso se busca una solución periódica con período 2π de la ecuación (12), que satisfaga una de las condiciones de contorno de primero, segundo o tercer género sobre los lados $r = l_1$ y $r = L_1$. Mostremos el tipo de una condición

de contorno de tercer género sobre el lado interior del anillo

$$\bar{k}_{11}\frac{\partial u}{\partial r}+\frac{\bar{k}_{12}}{r}\frac{\partial u}{\partial \varphi}=\varkappa_1^- u-g_1^-(\varphi), \quad r=l_1, \tag{24}$$

donde $\varkappa_1^-(\varphi)\geqslant 0$.

Si están prefijadas las condiciones de contorno de segundo género (21), (24) con $\varkappa_1^{\pm}(\varphi)\equiv 0$, entonces la solución del problema planteado existe, si se cumple la condición

$$\int_0^{2\pi}\int_{l_1}^{L_1} rf(r,\varphi)\,dr\,d\varphi+\int_0^{2\pi}[L_1 g_1^+(\varphi)+l_1 g_1^-(\varphi)]\,d\varphi=0. \tag{25}$$

En este caso la solución está determinada con exactitud hasta una constante.

Si *la región es un sector anular* ($l_1>0$, $L_2-l_2<2\pi$), entonces se plantea el problema de encontrar la solución de la ecuación (12), que satisfaga una de las condiciones de primero, segundo o tercer género sobre los lados del rectángulo G, en particular las condiciones (21), (24) para $r=L_1$ y $r=l_1$ y las condiciones de contorno de tercer género

$$\begin{aligned}
&\bar{k}_{21}\frac{\partial u}{\partial r}+\frac{\bar{k}_{22}}{r}\frac{\partial u}{\partial \varphi}=\varkappa_2^- u-g_2^-(r), \quad \varphi=l_2,\\
&-\bar{k}_{21}\frac{\partial u}{\partial r}-\frac{\bar{k}_{22}}{r}\frac{\partial u}{\partial \varphi}=\varkappa_2^+ u-g_2^+(r), \quad \varphi=L_2,
\end{aligned} \tag{26}$$

para $\varphi=l_2$ y $\varphi=L_2$, $\varkappa_2^{\pm}(r)\geqslant 0$.

Si son dadas las condiciones de contorno de segundo género (21), (24), (26) con $\varkappa_1^{\pm}(\varphi)\equiv 0$ y $\varkappa_2^{\pm}(r)\equiv 0$, entonces la solución del problema existe, si se cumple la condición

$$\int_{l_2}^{L_2}\int_{l_1}^{L_1} rf(r,\varphi)\,dr\,d\varphi+\int_{l_2}^{L_2}(L_1 g_1^+ + l_1 g_1^-)\,d\varphi+$$

$$+\int_{l_1}^{L_1}(g_2^- + g_2^+)\,dr=0. \tag{27}$$

Con esto la solución no es única y se determina con exactitud hasta una constante.

§ 2. Resolución de problemas de diferencias en un sistema cilíndrico de coordenadas

1. Esquemas de diferencias sin derivadas mixtas en el caso de simetría axial. Examinemos los problemas de contorno para las ecuaciones elípticas que no contienen derivadas mixtas en un sistema cilíndrico de coordenadas para el caso de simetría axial.

En el rectángulo $\bar{G} = \{l_1 \leqslant r \leqslant L_1, \quad l_3 \leqslant z \leqslant L_3, l_1 \geqslant 0\}$ se exige hallar la solución de la ecuación

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rk_1\frac{\partial u}{\partial r}\right)+\frac{\partial}{\partial z}\left(k_3\frac{\partial u}{\partial z}\right)-qu=-f(r, z), \quad (r, z)\in G, \tag{1}$$

que satisfaga sobre la frontera del rectángulo $\bar{G}$ las siguientes condiciones de contorno:

1) sobre el lado $r = l_1, \quad l_3 \leqslant z \leqslant L_3,$

$$u(r, z) = g_1^-(z), \text{ si } l_1 > 0, \tag{2}$$

ó

$$k_1\frac{\partial u}{\partial r} = \varkappa_1^- u - g_1^-(z), \text{ si } l_1 > 0, \qquad \lim_{r\to 0} rk_1\frac{\partial u}{\partial r} = 0, \text{ si } l_1 = 0; \tag{3}$$

2) sobre el lado $r = L_1, \; l_3 \leqslant z \leqslant L_3,$

$$u(r, z) = g_1^+(z) \tag{4}$$

ó

$$-k_1\frac{\partial u}{\partial r} = \varkappa_1^+ u - g_1^+(z); \tag{5}$$

3) sobre el lado $z = l_3, \; l_1 \leqslant r \leqslant L_1,$

$$u(r, z) = g_3^-(r) \tag{6}$$

ó

$$k_3\frac{\partial u}{\partial z} = \varkappa_3^- u - g_3^-(r); \tag{7}$$

4) sobre el lado $z = L_3, \; l_1 \leqslant r \leqslant L_1,$

$$u(r, z) = g_3^+(r) \tag{8}$$

ó

$$-k_3\frac{\partial u}{\partial z} = \varkappa_3^+ u - g_3^+(r). \tag{9}$$

Se supone, que los coeficientes satisfacen las condiciones

$$k_1(r, z) \geqslant c_1 > 0, \quad k_3(r, z) \geqslant c_1 > 0,$$
$$q(r, z) \geqslant 0,$$
$$\varkappa_1^{\pm}(z) \geqslant 0, \quad \varkappa_2^{\pm}(r) \geqslant 0.$$

En el caso, cuando $q \equiv 0$ y $\varkappa_\alpha^{\pm} \equiv 0$ en las condiciones de contorno (3), (5), (7), (9) ó $l_1 = 0$ y están dadas las condiciones de contorno de segundo género (5), (7), (9), exigimos el cumplimiento de la condición de solubilidad (véase (18) § 1).

Examinaremos combinaciones cualesquiera de las condiciones de contorno (2)—(9). Construyamos los esquemas de diferencias, correspondientes a los problemas de contorno indicados.

Introduzcamos en la región $\bar{G}$ una red rectangular no uniforme arbitraria

$$\bar{\omega} = \{(r_i, z_k) \in \bar{G}, \quad r_i = r_{i-1} + h_1(i), \quad 1 \leqslant i \leqslant N_1,$$
$$r_0 = l_1, \quad r_{N_1} = L_1, \quad z_k = z_{k-1} + h_3(k),$$
$$1 \leqslant k \leqslant N_3, \quad z_0 = l_3, \quad z_{N_3} = L_3\},$$

definamos los pasos medios

$$\hbar_\alpha(m) = \begin{cases} 0{,}5h_\alpha(1), & m = 0, \\ 0{,}5[h_\alpha(m) + h_\alpha(m+1)], & 1 \leqslant m \leqslant N_\alpha - 1, \\ 0{,}5h_\alpha(N_\alpha), & m = N_\alpha, \quad \alpha = 1, 3, \end{cases}$$

y la función reticular de una variable

$$\rho(i) = r_i, \quad 1 \leqslant i \leqslant N_1, \quad \rho(0) = \begin{cases} {}^1\!/_4 h_1(1), & l_1 = 0, \\ l_1, & l_1 > 0. \end{cases}$$

En el caso más simple cuando los coeficientes k_1, k_3, q y f son continuos, determinaremos los coeficientes del esquema de diferencias mediante las fórmulas

$$a_1(i, k) = \bar{r}_i k_1(\bar{r}_i, z_k), \quad a_3(i, k) = k_3(r_i, \bar{z}_k),$$
$$d(i, k) = q(r_i, z_k), \quad \varphi(i, k) = f(r_i, z_k),$$

donde $\bar{r}_i = r_i - 0{,}5h_1(i)$, $\bar{z}_k = z_k - 0{,}5h_3(k)$.

Utilizando las notaciones introducidas, aproximemos (1) por la ecuación en diferencias

$$\frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} + (a_3 y_{\bar{z}})_{\hat{z}} - dy = -\varphi, \quad 1 \leqslant i \leqslant N_1 - 1,$$
$$1 \leqslant k \leqslant N_3 - 1. \quad (10)$$

Aproximemos exactamente las condiciones de contorno de primer género (2), (4), (6), (8):

$$y(0, k) = g_1^-(z_k), \quad 0 \leqslant k \leqslant N_3, \tag{11}$$
$$y(N_1, k) = g_1^+(z_k), \quad 0 \leqslant k \leqslant N_3, \tag{12}$$
$$y(i, 0) = g_3^-(r_i), \quad 0 \leqslant i \leqslant N_1, \tag{13}$$
$$y(i, N_3) = g_3^+(r_i) \quad 0 \leqslant i \leqslant N_1. \tag{14}$$

El análogo de diferencias de las condiciones de contorno (3) tiene la forma

$$\frac{a_1^{+1}}{\rho h_1} y_r + (a_3 y_{\bar{z}})_{\hat{z}} - \left(d + \frac{\varkappa_1^-}{h_1}\right) y = -\varphi - \frac{g_1^-}{h_1}, \quad i = 0, \tag{15}$$

donde $1 \leqslant k \leqslant N_3 - 1$ y $\varkappa_1^- = g_1^- = 0$, si $l_1 = 0$. Las condiciones de contorno (5), (7), (9) se aproximan de la siguiente manera:

$$-\frac{a_1}{\rho h_1} y_{\bar{r}} + (a_3 y_{\bar{z}})_{\hat{z}} - \left(d + \frac{\varkappa_1^+}{h_1}\right) y = -\varphi - \frac{g_1^+}{h_1}, \quad i = N_1, \tag{16}$$

donde $1 \leqslant k \leqslant N_3 - 1$,

$$\frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} + \frac{a_3^{+1}}{h_3} y_z - \left(d + \frac{\varkappa_3^-}{h_3}\right) y = -\varphi - \frac{g_3^-}{h_3}, \quad k = 0, \tag{17}$$

$$\frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} - \frac{a_3}{h_3} y_{\bar{z}} - \left(d + \frac{\varkappa_3^+}{h_3}\right) y = -\varphi - \frac{g_3^+}{h_3}, \quad k = N_3, \tag{18}$$

donde $1 \leqslant i \leqslant N_1 - 1$. Aquí han sido utilizadas las notaciones $a_1^{+1} = a_1(i + 1, k)$, $a_3^{+1} = a_3(i, k + 1)$.

Si sobre los lados del rectángulo $\bar{G}$ que se intersecan son prefijadas las condiciones de contorno de tercer género, entonces en los nodos angulares de la red $\bar{\omega}$ se ponen las condiciones de contorno

$$\frac{a_1^{+1}}{\rho h_1} y_r + \frac{a_3^{+1}}{h_3} y_z - \left(d + \frac{\varkappa_1^-}{h_1} + \frac{\varkappa_3^-}{h_3}\right) y = -\varphi - \frac{g_1^-}{h_1} - \frac{g_3^-}{h_3}, \tag{19}$$

$i = k = 0$,

$$-\frac{a_1}{\rho h_1} y_{\bar{r}} + \frac{a_3^{+1}}{h_3} y_z - \left(d + \frac{\varkappa_1^+}{h_1} + \frac{\varkappa_3^-}{h_3}\right) y = -\varphi - \frac{g_1^+}{h_1} - \frac{g_3^-}{h_3}, \tag{20}$$

$i = N_1, \quad k = 0$,

$$\frac{a_1^{+1}}{\rho h_1} y_r - \frac{a_3}{h_3} y_{\bar{z}} - \left(d + \frac{\varkappa_1^-}{h_1} + \frac{\varkappa_3^+}{h_3}\right) y = -\varphi - \frac{g_1^-}{h_1} - \frac{g_3^+}{h_3}, \tag{21}$$

$i = 0, \quad k = N_3$,

$$-\frac{a_1}{\rho h_1} y_{\bar{r}} - \frac{a_3}{h_3} y_{\bar{z}} - \left(d + \frac{\varkappa_1^+}{h_1} + \frac{\varkappa_3^+}{h_3}\right) y = -\varphi - \frac{g_1^+}{h_1} - \frac{g_3^+}{h_3},$$

$i = N_1, \quad k = N_3$. (22)

Como antes, si $l_1 = 0$, entonces en (19) y (21) se debe poner $\varkappa_1^- = g_1^- = 0$.

Notemos que el problema de diferencias (10), (15)—(22) con las condiciones de contorno de tercer género sobre cada lado del rectángulo $\bar{G}$ puede ser escrito del modo compacto

$$\begin{aligned} &\Lambda y = -f, \quad 0 \leqslant i \leqslant N_1, \quad 0 \leqslant k \leqslant N_3, \\ &\Lambda = \Lambda_1 + \Lambda_2, \quad f = \varphi + \varphi_1/\hbar_1 + \varphi_3/\hbar_3, \end{aligned} \tag{23}$$

donde

$$\varphi_1(i, k) = \begin{cases} g_1^-, & i = 0, \\ 0, & 1 \leqslant i \leqslant N_1 - 1, \\ g_1^+, & i = N_1, \end{cases} \qquad \varphi_3(i, k) = \begin{cases} g_3^-, & k = 0, \\ 0, & 1 \leqslant k \leqslant N_3 - 1, \\ g_3^+, & k = N_3, \end{cases} \tag{24}$$

y los operadores de diferencias Λ_1 y Λ_3 se dan por las fórmulas

$$\Lambda_1 y = \begin{cases} \frac{a_1^{+1}}{\rho\hbar_1} y_r - \left(d_1 + \frac{\varkappa_1^-}{\hbar_1}\right) y, & i = 0, \\ \frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} - d_1 y, & 1 \leqslant i \leqslant N_1 - 1, \\ -\frac{a_1}{\rho\hbar_1} y_{\bar{r}} - \left(d_1 + \frac{\varkappa_1^+}{\hbar_1}\right) y, & i = N_1, \quad 0 \leqslant k \leqslant N_3, \end{cases} \tag{25}$$

$$\Lambda_3 y = \begin{cases} \frac{a_3^{+1}}{\hbar_3} y_z - \left(d_3 + \frac{\varkappa_3^-}{\hbar_3}\right) y, & k = 0, \\ (a_3 y_{\bar{z}})_{\hat{z}} - d_3 y, & 1 \leqslant k \leqslant N_3 - 1, \\ -\frac{a_3}{\hbar_3} y_{\bar{z}} - \left(d_3 + \frac{\varkappa_3^+}{\hbar_3}\right) y, & k = N_3, \quad 0 \leqslant i \leqslant N_1. \end{cases} \tag{26}$$

Aquí $d_1 + d_3 = d$, $d_1 \geqslant 0$ y $d_3 \geqslant 0$.

Hallemos las condiciones de solubilidad del esquema de diferencias (23) en el caso $d \equiv 0$ y $\varkappa_\alpha^\pm \equiv 0$, $\alpha = 1, 2$.

En el espacio H de las funciones reticulares, dadas sobre $\bar{\omega}$, definamos un producto escalar por la fórmula

$$(u, v) = \sum_{i=0}^{N_1} \sum_{k=0}^{N_3} u(i, k)\, v(i, k)\, \rho(i)\, \hbar_1(i)\, \hbar_3(k). \tag{27}$$

Definiremos los operadores A_1 y A_3, los cuales actúan en H, poniendo $A_\alpha = -\Lambda_\alpha$, $\alpha = 1, 3$. Entonces el esquema de

diferencias (23) se puede escribir en la forma de la ecuación operacional

$$Au = f, \quad A = A_1 + A_3. \tag{28}$$

Empleando la primera fórmula de Green de diferencias hallaremos, para el caso $d \equiv 0$ y $\varkappa_\alpha^\pm \equiv 0$, que

$$(Au, v) = \sum_{i=1}^{N_1} \sum_{k=0}^{N_2} h_1(i)\, h_3(k)\, (a_1 u_{\bar{r}} v_{\bar{r}})_{ik} + \\ + \sum_{i=0}^{N_1} \sum_{k=1}^{N_2} h_1(i)\, h_3(k)\, \rho(i)\, (a_3 u_{\bar{z}} v_{\bar{z}})_{ik} = (u, Av).$$

Por consiguiente, el operador A es autoconjugado y no negativo en H, al mismo tiempo $(Au, u) = 0$ únicamente en el caso, cuando $u(i, k) \equiv$ const ó $u(i, k) \equiv 0$. De aquí en virtud de la desigualdad de Cauchy-Buniakovski

$$(Au, u)^2 \leqslant (Au, Au)(u, u)$$

se deduce, que $Au = 0$ para $u \neq 0$, si u es constante sobre $\bar{\omega}$. De esta manera el núcleo del operador A consiste de las funciones reticulares, iguales a una constante sobre la red $\bar{\omega}$. Por eso el problema (28) es soluble, si se cumple la condición $(f, 1) = 0$ ó, en virtud de la definición de f, ó sea, la condición

$$\sum_{i=0}^{N_1} \sum_{k=0}^{N_2} \rho\varphi h_1 h_2 + \sum_{k=0}^{N_2} h_3 (\rho g_1^- + \rho g_1^+) + \sum_{i=0}^{N_1} h_1 \rho [g_3^- + g_3^+] = 0. \tag{29}$$

La condición (29) es el análogo de diferencias de la condición (18) de solubilidad del problema diferencial, correspondiente al problema de diferencias (23).

Si la condición (29) está cumplida, entonces la solución del problema (23) para $d \equiv 0$ y $\varkappa_\alpha^\pm \equiv 0$ existe, pero no es única y dos soluciones cualesquiera se diferencian en una constante. Por eso se puede separar una de las posibles soluciones, fijando el valor $y(i, k)$ en algún nodo de la red $\bar{\omega}$.

2. **Métodos directos.** Examinemos un caso, para el cual los problemas de diferencias (10)—(22) pueden ser resueltos por uno de los métodos directos expuestos en los capítulos III y IV.

Sean *los coeficientes* k_1, k_3 *y* q *de la ecuación* (1) *no dependientes de* z, es decir, $k_1 = k_1(r)$, $k_3 = k_3(r)$, $q = q(r)$, en las condiciones de contorno de tercer género (3), (5), los

coeficientes $\varkappa_1^+$ y $\varkappa_1^-$ son constantes, y en las condiciones (7), (9) $\varkappa_3^- = \varkappa_3^+ \equiv 0$.

Se admiten combinaciones cualesquiera de las condiciones de contorno (2)—(9). Se supone que la red $\bar{\omega}$ es uniforme respecto de z, es decir, $h_3(k) \equiv h_3$, y puede ser no uniforme respecto a r. Bajo las suposiciones indicadas los problemas de diferencias (10)—(22) pueden ser resueltos o bien mediante el método de reducción completa, o mediante el método combinado de reducción incompleta y de separación de variables.

Ilustremos la posibilidad de aplicación de los métodos directos con ayuda de un ejemplo, en el cual sobre los lados $r = l_1$ y $r = L_1$ se prefijan las condiciones de contorno de tercer (segundo) género (3), (5), y para $z = l_3$ y $z = L_3$, las de segundo género. Otras combinaciones de condiciones de contorno se examinan de manera análoga.

El esquema de diferencias, correspondiente al problema planteado, tiene la forma (23). En virtud de las suposiciones hechas más arriba los coeficientes del esquema de diferencias se determinan por las fórmulas (compárese con el punto 1) $a_1 = a_1(i) = \bar{r}_i k_1(\bar{r}_i)$, $a_3 = a_3(i) = k_3(r_i)$, $d = d(i) = q(r_i)$, así que $a_3^{+1} = a_3$. En la definición (25) del operador de diferencias Λ_1 elijamos $d_1 = d$, y en las fórmulas (26), que definen el operador Λ_3 pongamos $\varkappa_3^- = \varkappa_3^+ = 0$, $d_3 = 0$. Ya que la red $\bar{\omega}$ es uniforme respecto a z, entonces en (26) la expresión en diferencias $(a_3 y_{\bar{z}})\hat{z}$ se debe cambiar por $a_3 y_{\bar{z}z}$.

Reduzcamos ahora el problema de diferencias (23) a un sistema de ecuaciones vectoriales tripuntuales. Para esto introduzcamos el vector de las incógnitas

$$Y_k = (y(0, k), \quad y(1, k), \ldots, y(N_1, k)), \quad 0 \leqslant k \leqslant N_3,$$

el cual contiene los valores de la función reticular buscada sobre la k-ésima fila de la red $\bar{\omega}$, y el vector de los miembros derechos

$$F_k = (\theta_0 f(0, k), \quad \theta_1 f(1, k), \ldots, \theta_{N_1} f(N_1, k)),$$

$$0 \leqslant k \leqslant N_3,$$

donde $\theta_i = h_3^2/a_3(i)$, $0 \leqslant i \leqslant N_1$. Definamos la matriz cuadrada C, poniendo

$$CY_k = ((2E - \theta_0\Lambda_1)\, y(0, k), \ldots, (2E - \theta_{N_1}\Lambda_1) \times$$

$$\times\, y(N_1, k)).$$

Utilizando estas notaciones, escribamos el esquema de diferencias (23) en la forma vectorial

$$\begin{aligned} &CY_0 - 2Y_1 = F_0, \quad k = 0, \\ &-Y_{k-1} + CY_k - Y_{k+1} = F_k, \quad 1 \leqslant k \leqslant N_3 - 1, \\ &2Y_{N_3-1} + CY_{N_3} = F_{N_3}, \quad k = N_3. \end{aligned} \tag{30}$$

Para convencerse de esto, es suficiente multiplicar cada ecuación del esquema (23) por $(-\theta_l)$ y pasar a la escritura vectorial.

Recordemos, que el método de reducción completa para el sistema (30) fue construido en el punto 1 del § 4 del cap. III. El método combinado de reducción incompleta y de separación de variables fue examinado en el punto 2 del § 3 del cap. IV. Aquí la diferencia de los ejemplos examinados en los capítulos III y IV se encierra en otra definición del operador Λ_1. Pero como el operador de diferencias Λ_1 por anticipado es tripuntual, entonces esta diferencia no influye ni sobre la construcción de estos métodos, ni sobre el carácter de la dependencia del número de operaciones aritméticas del número de nodos de la red $\overline{\omega}$. Si $N_3 = 2^n$, entonces el número de operaciones aritméticas para los métodos indicados se estima por una magnitud $O(N_1 N_3 \log_2 N_3)$.

Como conclusión señalemos, que la aplicación del método combinado con la separación de una de las soluciones en el caso degenerado ($d \equiv 0$, $\varkappa_1^{\pm} = \varkappa_2^{\pm} \equiv 0$) está descrito detalladamente en el punto 2 del § 4 del cap. XII para el sistema cartesiano de coordenadas.

3. **Método de las direcciones variables.** Examinemos ahora un caso particular del problema (1)—(9), para el cual $k_1 = k_1(r)$, $k_3 = k_3(z)$, $q = \text{const}$, $\varkappa_\alpha^{\pm} = \text{const}$, $\alpha = 1, 3$, y sobre los lados del rectángulo $\overline{G}$ está dada cualquier combinación de las condiciones de contorno (2)—(9). En este caso se separan las variables en el problema (1)—(9).

Se supone, que la red $\overline{\omega}$ es arbitraria y no uniforme por cada dirección. Bajo las suposiciones hechas, los problemas de diferencias (10)—(22) pueden ser resueltos mediante el método de las direcciones variables con la elección óptima de los parámetros iterativos, el cual está citado en el capítulo XI para el caso de un sistema cartesiano de coordenadas.

Ilustremos la aplicación de este método en el ejemplo, en el cual sobre los lados del rectángulo $\overline{G}$ están prefijadas las condiciones de contorno de tercer género (3), (5), (7), (9).

El esquema de diferencias, correspondiente al problema (1), (3), (5), (7), (9), tiene la forma (23), donde los operadores Λ_1 y Λ_3 están definidos en (25), (26), y los coeficientes a_1, a_3, d_1 y d_3 se dan por las fórmulas $a_1(i) = \overline{r}_i k_1(\overline{r}_i)$, $a_3(k) = k_3(\overline{z}_h)$, $d_1 = d_3 = 0{,}5d$, $d = q$.

En el punto 1 fue mostrado, que el problema de diferencias (23) puede ser escrito en la forma de la ecuación operacional (28)

$$Au = f, \quad A = A_1 + A_3$$

en el espacio de Hilbert H de las funciones reticulares, definidas sobre $\overline{\omega}$. Indiquemos las propiedades fundamentales de los operadores A_1 y A_3:

1) los operadores A_1 y A_3 son conmutables, $A_1A_3 = A_3A_1$;

2) A_1 y A_3 son los operadores autoconjugados, $(A_\alpha u, v) = (u, A_\alpha v)$;

3) los operadores A_1 y A_3 son acotados y no negativos, es decir, para cualquier $u \in H$ están cumplidas las desigualdades

$$\delta_\alpha (u, u) \leqslant (A_\alpha u, u) \leqslant \Delta_\alpha (u, u),$$
$$\delta_\alpha \geqslant 0, \quad \Delta_\alpha \geqslant 0, \quad \alpha = 1, 3. \tag{31}$$

Realmente, la conmutatividad de los operadores A_1 y A_3 se deduce de la estructura de los operadores A_1 y A_3 y de la suposición con respecto a los coeficientes k_1, k_3, q y $\varkappa_\alpha^\pm$.

A continuación, utilizando la definición (27) del producto escalar en H y las fórmulas de Green de diferencias, obtendremos para A_1 y cualesquiera $u, v \in H$ la igualdad

$$(A_1u, v) = \sum_{i=1}^{N_1} \sum_{k=0}^{N_3} h_1(i)\, \hbar_3(k)\, (a_1 u_{\overline{r}} v_{\overline{r}})_{ih} + \\ + d_1(u, v) + \sum_{k=0}^{N_3} \hbar_3(k) [\varkappa_1^- \rho uv\,|_{i=0} + \varkappa_1^+ \rho uv\,|_{i=N_1}] \tag{32}$$

y la igualdad análoga para A_3

$$(A_3u, v) = \sum_{k=1}^{N_3} \sum_{i=0}^{N_1} \rho(i)\, \hbar_1(i)\, h_3(k) + (a_3 u_{\overline{z}} v_{\overline{z}})_{ih} + \\ + d_3(u, v) + \sum_{i=0}^{N_1} \rho(i)\, \hbar_1(i) [\varkappa_3^- uv\,|_{k=0} + \varkappa_3^+ uv\,|_{k=N_3}]. \tag{33}$$

Cambiando de lugares u y v, nos cercioramos de la autoconjugación de los operadores A_1 y A_3.

Si ponemos aquí $u = v$ y tenemos en cuenta las condiciones $k_1 \geqslant c_1 > 0$, $k_3 \geqslant c_1 > 0$, $q \geqslant 0$, $\varkappa_\alpha^{\pm} \geqslant 0$, $\alpha = 1, 3$, entonces hallaremos, que los operadores A_1 y A_3 son no negativos, es decir $(A_\alpha u, u) \geqslant 0$. Si está cumplida la condición

$$d_\alpha^2 + (\varkappa_\alpha^-)^2 + (\varkappa_\alpha^+)^2 \neq 0, \quad \alpha = 1, 3, \tag{34}$$

entonces la respectiva δ_α, es positiva. Supongamos que se cumple (34).

Daremos la estimación inferior de δ_α.

Del lema 16 del capítulo V para i fijo, $0 \leqslant i \leqslant N_1$, obtenemos la estimación

$$\delta_3 \sum_{k=0}^{N_3} h_3(k)\, u^2(i, k) \leqslant \sum_{k=1}^{N_3} h_3(k)\, a_3(k)\, u_{\bar z}^2(i, k) + \\ + d_3 \sum_{k=0}^{N_3} h_3(k)\, u^2(i, k) + \varkappa_3^- u^2(i, 0) + \varkappa_3^+ u^2(i, N_3), \tag{35}$$

donde $1/\delta_3 = \max\limits_{0 \leqslant k \leqslant N_3} v(k)$, y $v(k)$ es la solución del problema de contorno

$$(a_3 v_{\bar z})_{\hat z} - d_3 v = -1, \quad 1 \leqslant k \leqslant N_3 - 1,$$

$$\frac{a_3^{+1}}{h_3} v_z - \left(d_3 + \frac{\varkappa_3^-}{h_3}\right) v = -1, \quad k = 0, \tag{36}$$

$$-\frac{a_3}{h_3} v_{\bar z} - \left(d_3 + \frac{\varkappa_3^+}{h_3}\right) v = -1, \quad k = N_3.$$

Puesto que se cumple la condición (34), entonces la solución del problema (36) existe y es única. Multiplicando ahora (35) por $\rho(i)\, h_1(i)$ y sumando según i desde 0 hasta N_1, obtendremos la desigualdad $\delta_3(u, u) \leqslant (A_3 u, u)$. Resolviendo numéricamente el problema (36), determinaremos δ_3. Así pues, hemos hallado la constante δ_3. Análogamente se estima la constante δ_1: $1/\delta_1 = \max\limits_{0 \leqslant i \leqslant N_1} \bar v(i)$ donde $\bar v(i)$ es la solución del problema de contorno

$$\frac{1}{\rho}(a_1 \bar v_{\bar r})_{\hat r} - d_1 \bar v = -1, \quad 1 \leqslant i \leqslant N_1 - 1,$$

$$\frac{a_1^{+1}}{\rho h_1} \bar v_r - \left(d_1 + \frac{\varkappa_1^-}{h_1}\right) \bar v = -1, \quad i = 0, \tag{37}$$

$$-\frac{a_1}{\rho h_1} \bar v_{\bar r} - \left(d_1 + \frac{\varkappa_1^+}{h_1}\right) \bar v = -1, \quad i = N_1.$$

Obtengamos ahora estimados para Δ_1 y Δ_2. De (33) para $u = v$ hallaremos

$$(A_2 u,\ u) = \sum_{i=0}^{N_1} \rho(i)\, \hbar_1(i) \Big[\sum_{k=1}^{N_2} h_2(k)\, a_2(k)\, u_{\bar{z}}^2(i,\ k) + \\ + d_2 \sum_{k=0}^{N_2} \hbar_2(k)\, u^2(i,\ k) + \varkappa_2^- u^2(i,\ 0) + \varkappa_2^+ u^2(i,\ N_2) \Big].$$

Estimemos la expresión que se encuentra entre los corchetes. Del lema 16 del capítulo V obtenemos

$$d_2 \sum_{k=0}^{N_2} u^2(i,\ k)\, \hbar_2(k) + \varkappa_2^- u^2(i,\ 0) + \varkappa_2^+ u^2(i,\ N_2) \leqslant \\ \leqslant m_1 \Big[\sum_{k=1}^{N_2} a_2(k)\, u_{\bar{z}}^2(i,\ k)\, h_2(k) + \sum_{k=0}^{N_2} \hbar_2(k)\, u^2(i,\ k) \Big]. \quad (38)$$

donde $m_1 = \max\limits_{0 \leqslant k \leqslant N_2} w(k)$, y $w(k)$ es la solución del problema de contorno

$$\begin{aligned} &(a_2 w_{\bar{z}})_{\hat{z}} - w = -d_2, \quad 1 \leqslant k \leqslant N_2 - 1, \\ &\frac{a_2^{+1}}{\hbar_2} w_z - w = -\left(d_2 + \frac{\varkappa_2^-}{\hbar_2} \right), \quad k = 0, \\ &-\frac{a_2}{\hbar_2} w_{\bar{z}} - w = -\left(d_2 + \frac{\varkappa_2^+}{\hbar_2} \right), \quad k = N_2. \end{aligned} \quad (39)$$

Utilizando el lema 17 del capítulo V, tendremos

$$\sum_{k=1}^{N_2} a_2(k)\, u_{\bar{z}}^2(i,\ k)\, h_2(k) \leqslant m_2 \sum_{k=0}^{N_2} \hbar_2(k)\, u^2(i,\ k), \quad (40)$$

donde

$$m_2 = \max \left(\frac{a_2(N_2)}{\hbar_2^2(N_2)},\ \frac{a_2(1)}{\hbar_2^2(0)},\ \max_{1 \leqslant k \leqslant N_2 - 1} \frac{2}{\hbar_2(k)} \times \right. \\ \left. \times \left[\frac{a_2(k)}{h_2(k)} + \frac{a_2(k+1)}{h_2(k+1)} \right] \right).$$

De (38) y (40) se deduce la estimación

$$\sum_{k=1}^{N_2} h_2(k)\, a_2(k)\, u_{\bar{z}}^2(i,\ k) + d_2 \sum_{k=0}^{N_2} \hbar_2(k)\, u^2(i,\ k) + \\ + \varkappa_2^- u^2(i,\ 0) + \varkappa_2^+ u^2(i,\ N_2) \leqslant \Delta_2 \sum_{k=0}^{N_2} \hbar_2(k)\, u^2(i,\ k),$$

$\Delta_2 = m_1 + m_2(1 + m_1)$.

Multiplicando la desigualdad obtenida por $\rho(i)\, \hbar_1(i)$ y sumándola por i desde 0 hasta N_1, tendremos la estimación $(A_2 u,\ u) \leqslant \Delta_2 (u,\ u)$.

De forma análoga se encuentra Δ_1: $\Delta_1 = \bar{\bar{m}}_1 + \bar{\bar{m}}_2 (1 + \bar{\bar{m}}_1)$, donde $\bar{\bar{m}}_i = \max\limits_{0 \leqslant i \leqslant N_1} \bar{\bar{w}}(i)$, y $\bar{\bar{w}}(i)$ es la solución del problema de contorno

$$\begin{aligned}
&\frac{1}{\rho}(a_1 \bar{\bar{w}}_{\bar{r}})_{\hat{r}} - \bar{\bar{w}} = -d_1, \quad 1 \leqslant i \leqslant N_1 - 1, \\
&\frac{a_1^{+1}}{\rho \hbar_1} \bar{\bar{w}}_r - \bar{\bar{w}} = -\left(d_1 + \frac{\varkappa_1^-}{\hbar_1}\right), \quad i = 0, \\
&\frac{a_1}{\rho \hbar_1} \bar{\bar{w}}_{\bar{r}} - \bar{\bar{w}} = -\left(d_1 + \frac{\varkappa_1^+}{\hbar_1}\right), \quad i = N_1,
\end{aligned} \tag{41}$$

al mismo tiempo

$$\bar{\bar{m}}_2 = \max\left(\frac{a_1(N_1)}{\rho(N_1)\hbar_1^2(N_1)}, \frac{a_1(1)}{\rho(0)\hbar_1^2(0)}, \max_{1 \leqslant i \leqslant N_1 - 1} \frac{2}{\rho(i)\hbar_1(i)}\left[\frac{(a_1 i)}{\hbar_1(i)} + \frac{a_1(i+1)}{\hbar_1(i+1)}\right]\right).$$

Resolviendo numéricamente el problema (41), determinaremos $\bar{\bar{m}}_1$ y, por consiguiente, Δ_1. De esta forma son halladas las constantes δ_α y Δ_α, $\alpha = 1, 3$, que figuran en las desigualdades (31).

Recordemos que el esquema iterativo del método de las direcciones variables para la ecuación operacional (28) tiene la forma (véase el cap. XI)

$$\begin{aligned}
&B_{k+1}\frac{y_{k+1} - y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, \ y_0 \in H, \\
&B_k = (\omega_k^{(1)} E + A_1)(\omega_k^{(3)} E + A_3), \quad \tau_k = \omega_k^{(1)} + \omega_k^{(3)}.
\end{aligned} \tag{42}$$

En el punto 4 del § 1 del cap. XI para el esquema iterativo (42), cuyos operadores A_1 y A_3 satisfacen las propiedades 1)—3) enumeradas más arriba, fue construido el conjunto óptimo de los parámetros $\omega_k^{(1)}$ y $\omega_k^{(2)}$, $k = 1, 2, \ldots, n$. Utilizando este conjunto de parámetros la exactitud relativa $\varepsilon > 0$ ($\| y_n - u \|_D < \varepsilon \| y_0 - u \|_D$, $D = A, E$) se alcanza, si se cumplimentan $n \geqslant n_0(\varepsilon)$ iteraciones, donde

$$\begin{aligned}
&n_0(\varepsilon) = \frac{1}{\pi^2} \ln\frac{4}{\eta} \ln\frac{4}{\varepsilon}, \quad \eta = \frac{1-a}{1+a}, \\
&a = \sqrt{\frac{(\Delta_1 - \delta_1)(\Delta_3 - \delta_3)}{(\Delta_1 + \delta_3)(\Delta_3 + \delta_1)}}.
\end{aligned}$$

El conjunto de los parámetros óptimos $\omega_k^{(1)}$ y $\omega_k^{(2)}$ para el caso del segundo problema de contorno ($d = 0$, $\varkappa_\alpha^\pm \equiv 0$) fue construído en el punto 1 del § 4 del cap. XII.

4. Resolución de ecuaciones definidas sobre la superficie de un cilindro.

Examinemos ahora un método de resolución de análogos de diferencias de los problemas de contorno para una ecuación elíptica sin derivadas mixtas, definida sobre la superficie de un cilindro de radio R. Limitémonos al examen de la superficie cerrada según φ de un cilindro, puesto que los métodos de resolución de los problemas en el caso de una superficie no cerrada no se diferencian en nada de los métodos de resolución de los problemas planos en variables cartesianas.

Así pues, en la región $\bar{G} = \{l_2 \leqslant \varphi \leqslant L_2,\ l_3 \leqslant z \leqslant L_3,\ L_2 - l_2 = 2\pi\}$ se busca la solución de la ecuación

$$\frac{1}{R^2}\frac{\partial}{\partial\varphi}\left(k_2\frac{\partial u}{\partial\varphi}\right)+\frac{\partial}{\partial z}\left(k_3\frac{\partial u}{\partial z}\right)-qu=-f(\varphi, z),\quad (\varphi, z)\in G, \tag{43}$$

periódica respecto a φ con período 2π que satisfaga sobre los lados $z = l_3$ y $z = L_3$ o bien las condiciones de contorno de primer género $u(\varphi, z) = g_3^-(\varphi)$ para $z = l_3$, $u(\varphi, z) = g_3^+(\varphi)$ para $z = L_3$, ó de segundo o tercer género

$$\begin{aligned} &k_3\frac{\partial u}{\partial z}=\varkappa_3^- u-g_3^-(\varphi),\quad z=l_3\\ &-k_3\frac{\partial u}{\partial z}=\varkappa_3^+ u-g_3^+(\varphi),\quad z=L_3, \end{aligned} \tag{44}$$

ó de cualquiera de sus combinaciones. Se supone que los coeficientes satisfacen las condiciones

$$k_2(\varphi, z)\geqslant c_1>0,\quad k_3(\varphi, z)\geqslant c_1>0,$$
$$q(\varphi, z)\geqslant 0,\quad \varkappa_3^{\pm}(\varphi)\geqslant 0.$$

En la región $\bar{G}$ introduzcamos la red no uniforme

$$\bar{\omega}=\{(\varphi_j, z_k)\in\bar{G},\ \varphi_j=\varphi_{j-1}+h_2(j),\ 1\leqslant j\leqslant N_2,\ \varphi_0=l_2,$$
$$\varphi_{N_2}=L_2,\ z_k=z_{k-1}+h_3(k),\ 1\leqslant k\leqslant N_3,\ z_0=l_3,\ z_{N_3}=L_3\}$$

y definamos el paso medio

$$\hbar_2(j)=\begin{cases} 0{,}5\,[h_2(1)+h_2(N_2)], & j=0,\\ 0{,}5\,[h_2(j)+h_2(j+1)], & 1\leqslant j\leqslant N_2-1. \end{cases} \tag{45}$$

El paso medio $\hbar_3(k)$ está definido más arriba.

Aproximemos la ecuación (43) teniendo en cuenta la condición de periodicidad de la siguiente forma:

$$\begin{aligned} &(a_2 y_{\bar{\varphi}})_{\hat{\varphi}}+(a_3 y_{\bar{z}})_{\hat{z}}-dy=-\psi,\quad 0\leqslant j\leqslant N_2-1,\\ &1\leqslant k\leqslant N_3-1, \end{aligned} \tag{46}$$

donde son utilizadas las relaciones $y(j, k) = y(N_2 + j, k)$, $j = 0, -1$, $a_2(0, k) = a_2(N_2, k)$, $h_2(0) = h_2(N_2)$, las cuales son la consecuencia de la condición de periodicidad. En el caso de coeficientes suaves k_2, k_3, q y f los coeficientes en la ecuación (46) se pueden elegir, por ejemplo, así:

$$a_2(j, k) = \frac{1}{R^2} k_2(\varphi_j - 0{,}5h_2(j), z_k), \quad d(j, k) = q(\varphi_j, z_k),$$

$$a_3(j, k) = k_3(\varphi_j, z_k - 0{,}5h_3(k)), \quad \psi(j, k) = f(\varphi_j, z_k).$$

Las condiciones de contorno de primer género se aproximan exactamente

$$y(j, 0) = g_3^-(\varphi_j), \quad k = 0, \quad y(j, N_3) = g_3^+(\varphi_j), \quad k = N_3 \quad (47)$$

para $0 \leqslant j \leqslant N_2 - 1$, y el análogo de diferencias de las condiciones de contorno (44) de tercer género tiene la forma para $0 \leqslant j \leqslant N_2 - 1$:

$$\begin{aligned} &(a_2 y_{\bar{\varphi}})_{\hat{\varphi}} + \frac{a_3^{+1}}{h_3} y_z - \left(d + \frac{\varkappa_3^-}{h_3}\right) y = -\psi - \frac{g_3^-}{h_3}, \quad k = 0, \\ &(a_2 y_{\bar{\varphi}})_{\hat{\varphi}} - \frac{a_3}{h_3} y_{\bar{z}} - \left(d + \frac{\varkappa_3^+}{h_3}\right) y = -\psi - \frac{g_3^+}{h_3}, \quad k = N_3. \end{aligned} \quad (48)$$

En el problema (46), (47) las incógnitas son los valores $y(j, k)$ para $0 \leqslant j \leqslant N_2 - 1$, $1 \leqslant k \leqslant N_3 - 1$, y en el problema (46), (48) para los mismos valores j y $0 \leqslant k \leqslant N_3$.

Hallemos la condición de solubilidad del problema de diferencias (46), (48) en el caso, cuando $d \equiv 0$, $\varkappa_3^{\pm} \equiv 0$. Primeramente escribamos el esquema (46), (48) en la forma

$$\begin{aligned} &\Lambda y = -f, \quad 0 \leqslant j \leqslant N_2 - 1, \quad 0 \leqslant k \leqslant N_3, \\ &\Lambda = \Lambda_2 + \Lambda_3, \quad f = \psi + \psi_3/h_3, \end{aligned} \quad (49)$$

donde el operador en diferencias Λ_3 está definido en (26) con $d_3 = d$, el operador Λ_2 se da mediante la fórmula $\Lambda_2 y = (a_2 y_{\bar{\varphi}})_{\hat{\varphi}}$, $0 \leqslant j \leqslant N_2 - 1$, y

$$\psi_3(j, k) = \begin{cases} g_3^-(\varphi_j), & k = 0, \\ 0, & 1 \leqslant k \leqslant N_3 - 1, \\ g_3^+(\varphi_j), & k = N_3. \end{cases}$$

Sea ahora $d \equiv 0$ y $\varkappa_3^{\pm} \equiv 0$. Designemos mediante H el espacio de las funciones reticulares, definidas sobre $\bar{\omega}^* = \{(\varphi_j, z_k) \in \bar{\omega}, \quad 0 \leqslant j \leqslant N_2 - 1, \quad 0 \leqslant k \leqslant N_3\}$, en el

cual definiremos el producto escalar mediante la fórmula

$$(u, v) = \sum_{j=0}^{N_2-1} \sum_{k=0}^{N_3} u(j, k) v(j, k) \hbar_2(j) \hbar_3(k).$$

Definamos los operadores A_2 y A_3, los cuales actúan en H, con ayuda de las igualdades: $A_3 = -\Lambda_3$, $A_2 y = -\Lambda_2 \bar{y}$ donde $y(j, k) = \bar{y}(j, k)$ para $0 \leqslant j \leqslant N_2 - 1$, $0 \leqslant k \leqslant N_3$ y y satisface la condición de periodicidad $\bar{y}(j, k) = y(N_2 + j, k)$, $j = 0, -1$.

Empleando las notaciones introducidas, escribamos el esquema de diferencias (49) en la forma de la ecuación operacional

$$Au = f, \quad A = A_2 + A_3. \tag{50}$$

Tomando en consideración las condiciones de periodicidad, por medio de la fórmula de Green de diferencias obtendremos

$$(Au, v) = -(\Lambda \bar{u}, \bar{v}) = \sum_{k=0}^{N_3} \sum_{j=0}^{N_2-1} \hbar_3(k) h_2(j) (a_2 \bar{u}_{\bar{\varphi}} \bar{v}_{\bar{\varphi}})_{jh} +$$

$$+ \sum_{k=1}^{N_3} \sum_{j=0}^{N_2-1} \hbar_2(j) h_3(k) (a_3 \bar{u}_{\bar{z}} \bar{v}_{\bar{z}})_{jh} = (u, Av).$$

Por consiguiente, el operador A es autoconjugado en H. Además, examinando los valores de (Au, u), hallaremos, que el núcleo del operador A consiste de las funciones reticulares que toman valores constantes sobre la red ω^*. Por eso la solución del problema de diferencias (49) existe si está cumplida la condición $(f, 1) = 0$. Sustituyendo aquí f de (49), obtenemos

$$\sum_{j=0}^{N_2-1} \sum_{k=0}^{N_3} \hbar_2(j) \hbar_3(k) \psi(j, k) + \sum_{j=0}^{N_2-1} \hbar_2(j) [g_3^-(\varphi_j) + g_3^+(\varphi_j)] = 0.$$

Bajo el cumplimiento de esta condición la solución del problema de diferencias (46), (48) para $d \equiv 0$ y $\varkappa_3^{\pm} = 0$ existe y dos cualesquiera de estas soluciones se diferencian en una constante.

Examinemos el caso, cuando la solución de los problemas de diferencias (46)—(48) puede ser hallada por los métodos directos expuestos en los capítulos III y IV.

PRIMER CASO. Los coeficientes k_2, k_3 y q de la ecuación (43) dependen solamente de φ, $\varkappa_3^{\pm} = \text{const}$ y la red $\bar{\omega}$ es uniforme respecto de z. El problema de diferencias (46), (48) puede ser escrito en la forma del sistema de ecuaciones

vectoriales tripuntuales

$$\begin{aligned}&(C+2\alpha E)\,Y_0-2Y_1=F_0, \quad k=0,\\ &-Y_{k-1}+CY_k-Y_{k+1}=F_k, \quad 1\leqslant k\leqslant N_3-1,\\ &-2Y_{N_3-1}+(C+2\beta E)\,Y_{N_3}=F_{N_3}, \quad k=N_3,\end{aligned} \tag{51}$$

donde $N_3 = 2^n$, $n > 0$ es un número entero,

$Y_k = (y(0, k), y(1, k), \ldots, y(N_2-1, k))$,

$F_k = (\theta_0 f(0, k), \theta_1 f(1, k), \ldots, \theta_{N_2-1} f(N_2-1, k))$,

$CY_k = ((2E-\theta_0\Lambda_2)\,y(0, k), \ldots, (2E-\theta_{N_2-1}\Lambda_2)\,y(N_2-1, k))$

para $0 \leqslant k \leqslant N_3$. El operador Λ_2 está definido más arriba, $f(j, k)$ está dado en (49) y $\theta_j = h_3^2/a_3(j)$, $\alpha = h_3\varkappa_3^-$, $\beta = h_3\varkappa_3^+$.

Recordemos que en el punto 3 del § 4 del cap. III para resolver el problema (51) a la condición $\alpha^2 + \beta^2 \neq 0$ fue construido el método de reducción completa. Si $\alpha = \beta = 0$, pero $d \neq 0$, entonces el algoritmo del método está expuesto en el punto 1 del § 4 del cap. III. Para el último caso en el punto 2 del § 3 del cap. IV fue construido el método combinado de reducción incompleta y de separación de variables.

SEGUNDO CASO. Los coeficientes k_2, k_3 y q dependen sólo de z, $\varkappa_3^{\pm} = \text{const}$ y la red $\overline{\omega}$ es uniforme respecto a φ. El problema de diferencias (46), (48) se escribe en la forma del sistema de ecuaciones vectoriales tripuntuales

$$\begin{aligned}&-Y_{N_2-1}+CY_0-Y_1=F_0, \quad j=0,\\ &-Y_{j-1}+CY_j-Y_{j+1}=F_j, \quad 1\leqslant j\leqslant N_2-2,\\ &-Y_{N_2-2}+CY_{N_2-1}-Y_0=F_{N_2-1}, \quad j=N_2-1.\end{aligned} \tag{52}$$

Aquí $N_2 = 2^n$, $n > 0$ es un número entero.

$Y_j = (y(j, 0), y(j, 1), \ldots, y(j, N_3))$,

$F_j = (\theta_0 f(j, 0), \theta_1 f(j, 1), \theta_{N_3} f(j, N_3))$,

$CY_j = ((2E - \theta_0\Lambda_3)\,y(j, 0), \ldots, (2E - \theta_{N_3}\Lambda_3)\,y(j, N_3))$,

donde $0 \leqslant j \leqslant N_2 - 1$. El operador de diferencias Λ_3 está definido en (26) con $d_3 = d$ y $\theta_k = h_2^2/a_2(k)$, $0 \leqslant k \leqslant N_3$. El problema (52) puede ser resuelto por el método de reducción completa, construido en el punto 2 del § 4 del cap. III ó por el método combinado que utiliza el algoritmo de la transformación de Fourier discreta rápida de una función real periódica. Este algoritmo está construido en el punto 4 del § 1 del cap. IV.

En cada uno de los casos examinados los métodos directos se realizan con un gasto de $O(N_2N_3n)$ operaciones.

Como conclusión señalemos que si los coeficientes satisfacen las condiciones $k_2 = k_2(\varphi)$, $k_3 = k_3(z)$, $g = \text{const}$, $\varkappa_3^{\pm} = \text{const}$, y la red es no uniforme por cada dirección, entonces para encontrar la solución del problema (46), (48) se puede emplear el método de las direcciones variables con el conjunto óptimo de los parámetros:

$$B_{k+1}\frac{y_{k+1}-y_k}{\tau_{k+1}} + Ay_k = f, \quad k = 0, 1, \ldots, y_0 \in H,$$

$$B_k = (\omega_k^{(2)}E + A_2)(\omega_k^{(3)}E + A_3), \ \tau_k = \omega_k^{(2)} + \omega_k^{(3)}.$$

Aquí el operador $A_3 = -\Lambda_3$, $A_2 y = -\Lambda_2 \bar{y}$, donde el operador de diferencias Λ_3 está definido en (26) con $d_3 = 0{,}5d$, y $\Lambda_2 y = (a_2 y_{\bar{\varphi}})_{\hat{\varphi}} - 0{,}5dy$. Las constantes δ_α y Δ_α que son las cotas del operador A_α, se estiman de la forma siguiente: δ_3 y Δ_3 son halladas en el punto 3 del § 2, la constante δ_2 se encuentra exactamente: $\delta_2 = 0{,}5d$, y en calidad de Δ_2 se puede tomar

$$\Delta_2 = \max_{0 \leqslant j \leqslant N_2 - 1} \left[\frac{2}{h_2(j)} \left(\frac{a_2(j)}{h_2(j)} + \frac{a_2(j+1)}{h_2(j+1)} \right) + \frac{d}{2} \right].$$

§ 3. Resolución de problemas de diferencias en un sistema polar de coordenadas

1. Esquemas de diferencias para ecuaciones en un círculo y en un anillo. Examinemos los métodos de resolución de esquemas de diferencias para ecuaciones elípticas sin derivadas mixtas en un sistema polar de coordenadas. Primeramente estudiaremos el caso, cuando la región, donde se busca la solución, es un círculo o un anillo en el sistema cartesiano de coordenadas. En un sistema polar de coordenadas a las regiones indicadas les corresponde el rectángulo $\bar{G} = \{l_1 \leqslant r \leqslant L_1, \ l_2 \leqslant \varphi \leqslant L_2, \ l_1 \geqslant 0, \ L_2 - l_2 = 2\pi\}$. Se exige hallar la solución de la ecuación

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rk_1\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial}{\partial \varphi}\left(k_2\frac{\partial u}{\partial \varphi}\right) - qu = -f, \quad (r, \varphi) \in G, \qquad (1)$$

que sea periódica según φ con período 2π y que satisfaga sobre la frontera del rectángulo $\bar{G}$ las condiciones:

1) para $r = L_1$, $l_2 \leqslant \varphi \leqslant L_2$ ó bien las condiciones de contorno de primer género

$$u(r, \varphi) = g_1^{\dagger}(\varphi), \qquad (2)$$

ó de segundo o tercer género

$$-k_1 \frac{\partial u}{\partial r} = \varkappa_1^+ u - g_1^+ (\varphi); \tag{3}$$

2) para $r = l_1 > 0$, $l_2 \leqslant \varphi \leqslant L_2$ ó bien las condiciones de primer género

$$u (r, \varphi) = g_1^- (\varphi), \tag{4}$$

ó de segundo o tercer género

$$k_1 \frac{\partial u}{\partial r} = \varkappa_1^- u - g_1^- (\varphi); \tag{5}$$

para $r = l_1 = 0$ se pone la condición

$$\lim_{r \to 0} r k_1 \frac{\partial u}{\partial r} = 0, \tag{6}$$

la cual separa una solución acotada.

Se supone que los coeficientes satisfacen las condiciones $k_1 (r, \varphi) \geqslant c_1 > 0$, $k_2 (r, \varphi) \geqslant c_1 > 0$, $q (r, \varphi) \geqslant 0$, $\varkappa_1^{\pm} (\varphi) \geqslant 0$.

Examinaremos combinaciones cualesquiera de las condiciones de contorno (2)—(5). Construyamos los esquemas de diferencias correspondientes a los problemas de contorno indicados.

En la región $\bar{G}$ introduzcamos una red rectangular no uniforme arbitraria

$$\bar{\omega} = \{(r_i, \varphi_j) \in \bar{G}, \quad r_i = r_{i-1} + h_1 (i), \quad 1 \leqslant i \leqslant N_1,$$

$$r_0 = l_1, \quad r_{N_1} = L_1, \quad \varphi_j = \varphi_{j-1} + h_2 (j),$$

$$1 \leqslant j \leqslant N_2, \quad \varphi_0 = l_2, \quad \varphi_{N_2} = L_2\}.$$

El paso medio $\hbar_1 (i)$ está definido en el punto 1 del § 2, y el paso $\hbar_2 (j)$ en el punto 4 del § 2 por la fórmula (45). Definamos la función reticular $\rho (i)$:

$$\rho (i) = \begin{cases} l_1 + {}^1/_4 h_1 (1), & i = 0, \\ r_i + {}^1/_4 [h_1 (i+1) - h_1 (i)], & 1 \leqslant i \leqslant N_1 - 1, \\ L_1 - {}^1/_4 h_1 (N_1), & i = N_1. \end{cases} \tag{7}$$

En el caso más simple de coeficientes continuos k_1, k_2, q y f determinaremos los coeficientes del esquema de dife-

rencias por las fórmulas

$$a_1(i, j) = \overline{r}_1 k_1(\overline{r}_i, \varphi_j), \quad a_2(i, j) = k_2(r_i, \overline{\varphi}_j),$$
$$d(i, j) = q(r_i, \varphi_j), \quad \psi(i, j) = f(r_i, \varphi_j),$$

donde $\overline{r}_i = r_i - 0{,}5h_1(i)$, $\overline{\varphi}_j = \varphi_j - 0{,}5h_2(j)$.

Empleando las notaciones introducidas aproximemos (1) por la ecuación de diferencias

$$\Lambda y = \frac{1}{\rho}(a_1 y_{\overline{r}})_{\hat{r}} + \frac{1}{\rho^2}(a_2 y_{\overline{\varphi}})_{\hat{\varphi}} - dy = -\psi,$$
$$1 \leqslant i \leqslant N_1 - 1, \quad 0 \leqslant j \leqslant N_2 - 1. \tag{8}$$

Aquí para compacidad de la escritura son utilizadas las relaciones

$$y(i, j) = y(i, N_2 + j), \; j = 0, -1, \; a_2(i, 0) = a_2(i, N_2),$$
$$h_2(0) = h_2(N_2), \tag{9}$$

que son consecuencia de la condición de periodicidad.

Las condiciones de contorno (2), (4) se aproximan exactamente

$$y(N_1, j) = g_1^+(\varphi_j), \quad y(0, j) = g_1^-(\varphi_j),$$
$$0 \leqslant j \leqslant N_2 - 1. \tag{10}$$

El análogo de diferencias de las condiciones de contorno de tercer género (3), (5) tiene la forma (para $0 \leqslant j \leqslant N_2 - 1$)

$$\Lambda y = -\frac{a_1}{\rho \hbar_1} y_{\overline{r}} + \frac{1}{\rho^2}(a_2 y_{\overline{\varphi}})_{\hat{\varphi}} - \left(d + \frac{r\varkappa_1^+}{\rho \hbar_1}\right) y =$$
$$= -\psi - \frac{r g_1^+}{\rho \hbar_1}, \quad i = N_1, \tag{11}$$

$$\Lambda y = \frac{a_1^{+1}}{\rho \hbar_1} y_r + \frac{1}{\rho^2}(a_2 y_{\overline{\varphi}})_{\hat{\varphi}} - \left(d + \frac{r\varkappa_1^-}{\rho \hbar_1}\right) y =$$
$$= -\psi - \frac{r q_1^-}{\rho \hbar_1}, \quad i = 0. \tag{12}$$

Aquí se utilizan las relaciones (9).

Queda por construir una condición de contorno de diferencias sobre el lado $r = l_1$ para el caso, cuando $l_1 = 0$. Ya que todos los nodos, que están sobre el lado $r = 0$, se identifican, entonces

$$y(0, j) = y_0, \quad 0 \leqslant j \leqslant N_2 - 1. \tag{13}$$

A continuación, como el origen de coordenadas es un punto interior del círculo, entonces, escribiendo la ecuación (1)

en un sistema cartesiano de coordenadas y aproximándola sobre una red radial-anular a la condición (6), obtendremos

$$\Lambda y = \frac{1}{2\pi\rho\hbar_1} \sum_{j=0}^{N_2-1} a_1^{+1} y_r \hbar_2 - dy = -\psi, \quad i = 0,$$
$$d(0, j) = d_0, \ \psi(0, j) = \psi_0, \quad 0 \leqslant j \leqslant N_2 - 1. \tag{14}$$

Aquí y_0, d_0 y ψ_0 son los valores de las respectivas funciones reticulares en el centro del círculo.

Así pues, en el caso del círculo tenemos la condición de contorno no local (13), (14) sobre el lado $r = 0$ del rectángulo $\bar{G}$. Los esquemas de diferencias están construidos.

Para la aproximación de diferencias de la ecuación (1) en torno de $r = 0$ se utiliza a menudo otra red respecto a r, en la cual no está contenido el punto $r = 0$:

$$\bar{\omega} = \{(r_i, \varphi_j) \in \bar{G}, \ r_i = (i + 0{,}5)\, h_1, \quad 0 \leqslant i \leqslant N_1, \ r_{N_1} = L_1,$$
$$\varphi_j = \varphi_{j-1} + h_2(j), \quad 1 \leqslant j \leqslant N_2, \quad \varphi_0 = l_2, \quad \varphi_{N_2} = L_2\}$$

(para simplificar suponemos que la red es uniforme según r).

Entonces $a_1(i, j) = \bar{r}_i k_1(\bar{r}_i, \varphi_j)$, $a_2(i, j) = k_2(r_i, \varphi_j)$ etc., donde $r_i = ih_1$. Las ecuaciones (8) se mantienen sin cambios, y para $i = 0$ se escribe la siguiente ecuación en diferencias:

$$\Lambda y = \frac{\bar{r}_1}{r_0 h_1} a_1(1, j)\, y_r(1, j) + \frac{1}{r_0^2} (a_2 y_{\bar{\varphi}})_{\hat{\varphi}} - dy = -\psi$$

(aquí $r_0 = 0{,}5h_1$, $\bar{r}_1 = h_1$), la cual es un análogo de la condición de contorno de tercer género.

La condición para $r = 0$ está ausente y no se puede determinar el valor de y para $r = 0$ de las ecuaciones en diferencias escritas.

2. Solubilidad de los problemas de contorno de diferencias. En el punto 1 fueron construidos esquemas de diferencias que aproximan a los problemas (1)—(6). Para un círculo el esquema está prefijado por las fórmulas (8), (10), (11), (14) y para un anillo, por las fórmulas (8), (10), (12). Investiguemos la pregunta sobre la solubilidad de los esquemas indicados.

Designemos mediante $\bar{\omega}^*$ la parte de la red $\bar{\omega}$: $\bar{\omega}^* = \{(r_i, \varphi_j) \in \bar{\omega}, \quad 0 \leqslant j \leqslant N_1, \quad 0 \leqslant j \leqslant N_2 - 1\}$. El espacio H consiste de las funciones reticulares definidas sobre $\bar{\omega}^*$ y que satisfacen la condición complementaria $y(0, j) = \text{const}$, $0 \leqslant j \leqslant N_2 - 1$, si $l_1 = 0$. Definiremos

el producto escalar en H por la fórmula

$$(u, v) = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2-1} u(i, j) v(i, j) \rho(i) \hbar_1(i) \hbar_2(j).$$

Se puede mostrar, que si la función $\rho(i)$ está definida por la fórmula (7), entonces es cierta la igualdad

$$(1, 1) = 0{,}5 (L_1^2 - l_1^2)(L_2 - l_2) = \pi (L_2^2 - l_1^2), \qquad (15)$$

es decir, el cuadrado de la norma de la función, idénticamente igual a la unidad sobre $\bar{\omega}^*$, es igual al área del círculo ($l_1 = 0$) o del anillo ($l_1 > 0$). Además si la región a examinar es un círculo, entonces utilizando la constancia respecto de j para $i = 0$ de las funciones reticulares de H y la igualdad $\sum_{j=0}^{N_2-1} \hbar_2(j) = L_2 - l_2 = 2\pi$, se puede obtener la siguiente expresión para el producto escalar introducido más arriba:

$$(u, v) = \rho(0) \hbar_1(0) 2\pi u_0 v_0 + \sum_{i=1}^{N_1} \sum_{j=0}^{N_2-1} u(i, j) v(i, j) \rho(i) \hbar_1(i) \hbar_2(j), \qquad (16)$$

donde $u_0 = u(0, j)$, $v_0 = v(0, j)$.

Investiguemos la solubilidad de los problemas de diferencias (8), (11), (13) y (14) para $l_1 = 0$ y (8), (11), (12) para $l_1 > 0$, si $d \equiv 0$, $\varkappa_1^+ = \varkappa_1^- \equiv 0$. Escribamos los problemas de diferencias indicados mas arriba en la forma de la ecuación operacional

$$Au = f, \qquad (17)$$

donde el operador A se determina de la siguiente forma: $Ay = -\Lambda\bar{y}$, $\underline{y}(l, j) = \bar{y}(i, j)$ para $0 \leqslant i \leqslant N_1$, $0 \leqslant j \leqslant N_2 - 1$ y $\bar{y}$ satisface las condiciones de periodicidad (9), además $y(0, j) = \bar{y}(0, j) = \text{const}$.

Examinemos primeramente el operador A, correspondiente al operador de diferencias Λ del problema (8), (11), (13), (14). Tomando en consideración que la primera fórmula de Green de diferencias para las funciones que satisfacen la condición de periodicidad (9), obtiene la forma

$$\sum_{j=0}^{N_2-1} (a_2 u_{\bar{\varphi}})_{\hat{\varphi}} v \hbar_2 = -\sum_{j=0}^{N_2-1} a_2 u_{\bar{\varphi}} v_{\bar{\varphi}} h_2,$$

tendremos, considerando (16)

$$(Au,\ v) = -(\Lambda\bar{u},\ \bar{v}) = \sum_{j=0}^{N_2-1} \hbar_2 \Big(\sum_{i=1}^{N_1} h_1 a_1 \bar{u}_{\bar{r}} \bar{v}_{\bar{r}} + \sum_{i=0}^{N_1} \rho \hbar_1 d\bar{u}\bar{v} + r\varkappa_1^+ \bar{u}\bar{v}|_{i=N_1} \Big) + \sum_{i=1}^{N_1} \frac{\hbar_1}{\rho} \sum_{j=0}^{N_2-1} h_2 a_2 \bar{u}_{\bar{\varphi}} \bar{v}_{\bar{\varphi}} = -(\bar{u},\ \Lambda\bar{v}) = (u,\ Av).$$

Por consiguiente, el operador A es autoconjugado en H.

Para el operador A, correspondiente al operador de diferencias Λ del problema (8), (11), (12), obtendremos la igualdad análoga

$$(Au,\ v) = \sum_{j=0}^{N_2-1} \hbar_2 \Big(\sum_{i=1}^{N_1} h_1 a_1 \bar{u}_{\bar{r}} \bar{v}_{\bar{r}} + \sum_{i=0}^{N_1} \rho \hbar_1 d\bar{u}\bar{v} + r\varkappa_1^- \bar{u}\bar{v}|_{i=0} + r\varkappa_1^+ \bar{u}\bar{v}|_{i=N_1} \Big) + \sum_{i=0}^{N_1} \frac{\hbar_1}{\rho} \sum_{j=0}^{N_2-1} h_2 a_2 \bar{u}_{\bar{\varphi}} \bar{v}_{\bar{\varphi}} = (u,\ Av),$$

de la cual se deduce la autoconjugación del operador A.

Si $d \equiv 0$ y $\varkappa_i^{\pm} \equiv 0$, entonces de la autoconjugación del operador A y de la desigualdad de Cauchy-Buniakovsky $(Au,\ u) \leqslant \| Au \| \| u \|$ se deduce que el núcleo del operador A consiste de las funciones reticulares, iguales a una constante sobre la red $\bar{\omega}^*$. Por eso la condición de existencia de solución de la ecuación operacional (17) tiene la forma $(f,\ 1) = 0$. Para el problema (8), (11), (13), (14) a ella le corresponde la condición

$$\sum_{j=0}^{N_1} \sum_{j=0}^{N_2-1} \psi(i,\ j)\, \rho(i)\, \hbar_1(i)\, \hbar_2(j) + L_1 \sum_{j=0}^{N_2-1} \hbar_2(j)\, g_1^+(\varphi_j) = 0. \qquad (18)$$

la cual es un análogo de diferencias de la condición (23) del § 1. Para el problema (8), (11), (12) la condición de solu-

bilidad tiene la forma

$$\sum_{i=0}^{N_1}\sum_{j=0}^{N_2-1}\psi(i,j)\,\rho(i)\,\hbar_1(i)\,\hbar_2(j)+$$

$$+\sum_{j=0}^{N_2-1}\hbar_2(j)\,[L_1 g_1^{+}(\varphi_j)+l_1 g_1^{-}(\varphi_j)]=0$$

y es el análogo de la condición (25) del § 1, la cual asegura la solubilidad del problema diferencial correspondiente para un anillo.

Si se cumplen las condiciones indicadas, entonces las soluciones de los problemas examinados existen y dos soluciones cualesquiera se diferencian en una constante. La solución normal de estos problemas satisface la condición $(\bar{y}, 1) = 0$.

Sea y una de las soluciones, la cual se puede hallar, por ejemplo, fijando la solución buscada en un nodo de la red. Entonces, teniendo en cuenta la igualdad (15), obtendremos, que la función

$$\bar{y}=y-\frac{(y,1)}{\pi(L_1^2-l_1^2)}=y-\frac{(y,1)}{(1,1)}$$

es la solución normal.

OBSERVACION. Si determinamos la función reticular $\rho(i)$ mediante las fórmulas

$$\rho(i)=r_i,\ 1\leqslant i\leqslant N_1,\ \rho(0)=\begin{cases}h_1(0)/4, & l_1=0,\\ l_1, & l_1>0,\end{cases}$$

entonces cambia solamente la igualdad (15) para el caso, cuando $l_1 = 0$. En este caso tendremos

$$(1,1)=\pi L_1^2+\frac{h_1^2(1)}{4}\pi=\pi\left(L_1^2+\frac{h_1^2(1)}{4}\right).$$

3. Principio de superposición para un problema en el círculo. La resolución de los problemas de diferencias en un círculo se complica por la presencia de la condición de contorno no local (14), prefijado para $i = 0$. Observemos, que si el problema es degenerado y se cumple la condición de solubilidad (18), entonces es cómodo separar una de las soluciones, fijando su valor en el centro del círculo, es decir, prefijando $y(0, j) = y_0$, $0 \leqslant j \leqslant N_2 - 1$. En este caso la condición (14) no se utiliza, y el problema obtenido con y_0 dado es análogo al problema planteado para el anillo con una condición de contorno de primer género sobre la circun-

ferencia interior. Sea ahora el problema de diferencias (8), (11), (13), (14) no degenerado. Mostremos que su solución se puede hallar, resolviendo dos problemas auxiliares con condiciones de contorno locales de primer género para $i = 0$, $0 \leqslant j \leqslant N_2 - 1$.

Buscaremos la solución del problema (8), (11), (13), (14) en la forma

$$y(i, j) = v(i, j) + y_0 w(i, j), \quad 0 \leqslant i \leqslant N_1, \quad 0 \leqslant j \leqslant N_2 - 1, \tag{19}$$

donde y_0 es el valor de la solución buscada en el centro del círculo, y $v(i, j)$ y $w(i, j)$ satisfacen las condiciones de periodicidad

$$v(i, j) = v(i, N_2 + j), \quad w(i, j) = w(i, N_2 + j), \quad = 0, -1$$

y son las soluciones de las siguientes problemas de contorno:

$$\left.\begin{aligned} &\frac{1}{\rho}(a_1 v_{\bar{r}})_{\hat{r}} + \frac{1}{\rho^2}(a_2 v_{\bar{\varphi}})_{\hat{\varphi}} - dv = -\psi, \quad 1 \leqslant i \leqslant N_1 - 1, \\ &v(0, j) = 0, \quad i = 0, \qquad\qquad 0 \leqslant j \leqslant N_2 - 1, \\ &-\frac{a_1}{\rho \hbar_1} v_{\bar{r}} + \frac{1}{\rho^2}(a_2 v_{\bar{\varphi}})_{\hat{\varphi}} - \left(d + \frac{r\varkappa_1^+}{\rho \hbar_1}\right) v = -\psi - \frac{r g_1^+}{\rho \hbar_1}, \quad i = N_1, \end{aligned}\right\} \tag{20}$$

$$\left.\begin{aligned} &\Lambda w = \frac{1}{\rho}(a_1 w_{\bar{r}})_{\hat{r}} + \frac{1}{\rho^2}(a_2 w_{\bar{\varphi}})_{\hat{\varphi}} - dw = 0, \\ &1 \leqslant i \leqslant N_1 - 1, \quad 0 \leqslant j \leqslant N_2 - 1, \\ &w(0, j) = 1, \quad i = 0, \\ &-\frac{a_1}{\rho \hbar_1} w_{\bar{r}} + \frac{1}{\rho^2}(a_2 w_{\bar{\varphi}})_{\hat{\varphi}} - \left(d + \frac{r\varkappa_1^+}{\rho \hbar_1}\right) w = 0, \\ &i = N_1. \end{aligned}\right\} \tag{21}$$

Es evidente, que la función y, definida de acuerdo con (19), satisface la ecuación (8) y las condiciones (11), (13). Queda por determinar y_0. Sustituyendo (19) en la condición (14) aún no utilizada y teniendo en cuenta las condiciones de contorno para v y w, obtenemos

$$y_0 = \frac{\left[2\pi\rho\hbar_1\psi_0 + \sum\limits_{j=0}^{N_2-1} a_1^+ v_{\bar{r}} \hbar_2(j)\right]_{i=0}}{\left[2\pi\rho\hbar_1 d_0 - \sum\limits_{j=0}^{N_2-1} a_1^+ w_{\bar{r}} \hbar_2(j)\right]_{i=0}}. \tag{22}$$

Mostremos, que el denominador en (22) es diferente de cero. Para esto multipliquemos la ecuación (21) escalarmente por w. Utilizando las condiciones de contorno para w, la relación de periodicidad y las fórmulas de Green de diferencias, obtendremos

$$0 = \sum_{j=1}^{N_1-1} \sum_{j=0}^{N_2-1} (\Lambda w)\, w \rho \hbar_1 \hbar_2 =$$

$$= -\sum_{j=0}^{N_2-1} \hbar_2 \left(a^{\dagger 1} w_r|_{i=0} + L_1 \varkappa_1^{\dagger} w^2|_{i=N_1}\right) -$$

$$-\sum_{i=1}^{N_1} \sum_{j=0}^{N_2-1} a_1 w_{\bar{r}}^2 h_1 \hbar_2 - \sum_{i=0}^{N_1} \sum_{j=0}^{N_2-1} \hbar_1 \left[\frac{h_2}{\rho} a_2 w_{\bar{\varphi}}^2 + \hbar_2 \rho\, dw^2\right].$$

Ya que la función w es diferente de una constante, $d \geqslant 0$, $a_\alpha \geqslant c_1 > 0$, $\alpha = 1, 2$, y $\varkappa_1^{\dagger} \geqslant 0$, al mismo tiempo $d^2 + (\varkappa^{\dagger})^2 \neq 0$, entonces de aquí obtenemos, que

$$\sum_{j=0}^{N_2-1} a^{\dagger 1} w_r \hbar_2|_{i=0} < 0$$

y, por consiguiente, el denominador en la fórmula (22) es diferente de cero.

Así pues, la solución del problema inicial (8), (11), (13), (14) se reduce a la resolución de los dos problemas (20) y (21) con condiciones de contorno locales y a encontrar y_0 por la fórmula (22). La solución y buscada se encuentra por la fórmula (19).

Señalemos, que si sobre el lado $r = L_1$ está dada la condición de contorno de primer género $y(N_1, j) = g_1^{\dagger}(\varphi_j)$, entonces para las funciones v y w en lugar de las condiciones de tercer género en (20) y (21) se debe poner $v(N_1, j) = g_1^{\dagger}(\varphi_j)$ y $w(N_1, j) = 0$ para $0 \leqslant j \leqslant N_2 - 1$. La fórmula (22) para y_0 se conserva. Si los coeficientes k_1, k_2, q y $\varkappa_1^{\dagger}$ no dependen de φ, entonces tampoco depende de φ la solución w del problema (21). En este caso para las funciones w nosotros tenemos el problema de una dimensión

$$\frac{1}{\rho}(a_1 w_{\bar{r}})_{\hat{r}} - dw = 0, \quad 1 \leqslant i \leqslant N_1 - 1,$$

$$w(0, j) = 1, \quad i = 0,$$

$$-\frac{a_1}{\rho \hbar_1} w_r - \left(d + \frac{r \varkappa_1^{\dagger}}{\rho \hbar_1}\right) w = 0, \quad i = N_1,$$

el cual se resuelve por el método de factorización.

4. Métodos directos de resolución de ecuaciones en un círculo y en un anillo. De lo dicho más arriba se deduce, que es suficiente limitarse al examen de los métodos de resolución de los problemas de diferencias (8), (10)—(12). Al principio examinemos el caso, para el cual los problemas de diferencias indicados pueden ser resueltos por uno de los métodos directos expuestos en los capítulos III y IV.

Sean los coeficientes k_1, k_2 y q de la ecuación (1) no dependientes de φ: $k_1 = k_1(r)$, $k_2 = k_2(r)$, $q = q(r)$. Tal situación tiene lugar para la ecuación de Poisson en un sistema polar de coordenadas. Sean, además, $\varkappa_1^-$ y $\varkappa_1^+$ constantes en las condiciones de contorno de tercer género (11), (12). Se supone que la red ω es uniforme según φ, es decir, $h_2(j) \equiv h_2$, y puede ser no uniforme según r. Bajo las suposiciones indicadas las ecuación en diferencias (8) con cualquier combinación de las condiciones de contorno (11)—(12) puede ser resuelta ó bien por el método de reducción completa, ó por el método combinado de reducción incompleta y el de separación de variables.

Ilustremos la posibilidad de aplicación de los métodos directos en un ejemplo, en el cual sobre los lados $r = l_1$ y $r = L_1$ están prefijadas las condiciones de contorno de tercer (segundo) género (11), (12). Otras combinaciones de condiciones de contorno se examinan análogamente.

En virtud de las suposiciones hechas los coeficientes del esquema de diferencias se determinan mediante las fórmulas

$$a_1(i) = \bar{r}_i k_1(\bar{r}_i),\ a_2(i) = k_2^{\mathrm{II}}(r_i),$$
$$d(i) = q(r_i),$$

y como la red $\bar{\omega}$ es uniforme según φ, entonces el operador de diferencias $(a_2 y_{\bar{\varphi}})_{\hat{\varphi}}$ se cambia por $a_2 y_{\bar{\varphi}\varphi}$.

Reduzcamos el problema de diferencias (8), (11), (12) al sistema de ecuaciones vectoriales tripuntuales

$$\begin{aligned} &-Y_{N_2-1} + CY_0 - Y_1 = F_0, \quad j = 0, \\ &-Y_{j-1} + CY_j - Y_{j+1} = F_j, \quad 1 \leqslant j \leqslant N_2 - 2, \\ &-Y_{N_2-2} + CY_{N_2-1} - Y_0 = F_{N_2-1}, \quad j = N_2 - 1. \end{aligned} \tag{23}$$

Aquí para $0 \leqslant j \leqslant N_2 - 1$ son utilizadas las notaciones:

$$Y_j = (y(0, j), y(1, j), \ldots, y(N_1, j)),$$
$$F_j = (0_0 f(0, j), 0_1 f(1, j), \ldots, 0_{N_1} f(N_1, j)),$$
$$CY_j = ((2E - 0_0 \Lambda_1) y(0, j), \ldots, (2E - 0_{N_1} \Lambda_1) y(N_1, j)),$$

donde

$$f(i,\ j)=\begin{cases}\psi(0,\ j)+\dfrac{l_1 g_1^-(\varphi j)}{\rho(0)\,h_1(0)}, & i=0,\\ \psi(i,\ j), & 1\leqslant i\leqslant N_1-1,\\ \psi(N_1,\ j)+\dfrac{L_1 g_1^+(\varphi j)}{\rho(N_1)\,h(N_1)}, & i=N_1,\end{cases}\tag{24}$$

el operador en diferencias Λ_1 actúa de la siguiente forma:

$$\Lambda_1 y=\begin{cases}\dfrac{a_1^{+1}}{\rho h_1}y_r-\left(d+\dfrac{r\varkappa_1^-}{\rho h_1}\right)y, & i=0,\\ \dfrac{1}{\rho}(a_1 y_{\bar r})_{\mathring r}-dy, & 1\leqslant i\leqslant N_1-1,\\ -\dfrac{a_1}{\rho h_1}y_{\bar r}-\left(d+\dfrac{r\varkappa_1^+}{\rho h_1}\right)y, & i=N_1,\end{cases}\tag{25}$$

y, finalmonto, $\theta_i=\rho^2(i)\,h_2^2/a_2(i)$, $0\leqslant i\leqslant N_1$.

El sistema (23) se obtiene de (8), (11), y (12) al multiplicar cada ecuación por el correspondiente θ_i y pasar a la escritura vectorial.

Recordemos que el algoritmo del método de reducción completa para el sistema (23) está descrito en el punto 2 del § 4 del cap. III. En el método combinado se utiliza el algoritmo de la transformación de Fourier discreta rápida, el cual está citado en el punto 4 del § 1 del cap. IV. Estos métodos se caracterizan por la estimación $O(N_1N_2\log_2 N_2)$ de operaciones aritméticas para $N_2=2^n$.

5. Método de las direcciones variables. Supongamos ahora que los coeficientes en la ecuación (1) y en las condiciones de contorno (3), (5) satisfacen las condiciones $k_1=k_1(r)$, $k_2=k_2(\varphi)$, $q=\text{const}$, $\varkappa_i^{\pm}=\text{const}$, es decir, para el problema (1), (3), (5) aplicaremos el método de separación de variables. Se supone que la red $\bar\omega$ es no uniforme por cada dirección. Examinemos la ecuación en diferencias (8) con cualquier combinación de las condiciones de contorno (10)—(12). Bajo las suposiciones hechas se separan las variables en el esquema de diferencias, y su solución aproximada puede ser hallada con ayuda del método de las direcciones variables con la colección óptima de los parámetros iterativos.

Para un ejemplo examinaremos el problema (8), (11), (12) con condiciones de contorno de tercer género para $r=l_1$ y $r=L_1$. Escribamos este problema en la forma

$$\begin{aligned}&\bar\Lambda y=-\bar f,\quad 0\leqslant i\leqslant N_1,\quad 0\leqslant j\leqslant N_2-1,\\ &\bar\Lambda=\bar\Lambda_1+\bar\Lambda_2,\quad \bar f=\rho^2 f,\end{aligned}\tag{26}$$

donde $\tilde{\Lambda}_1 = \rho^2 \Lambda_1$, el operador Λ_1 está definido en (25), el operador $\bar{\Lambda}_2$ se da mediante la igualdad $\bar{\Lambda}_2 y = (a_2 y_{\bar{\varphi}})_{\hat{\varphi}}$, al mismo tiempo se cumplen las relaciones (9), y el segundo miembro f está definido en (24). La ecuación (26) es obtenida de (8), (11) y (12) al multiplicar por ρ^2.

En virtud de las suposiciones hechas los coeficientes del esquema de diferencias (26) se eligen por las fórmulas

$$a_1(i) = \bar{r}_i k_1(\bar{r}_i), \quad a_2(j) = k_2^r(\bar{\varphi}_j), \quad d = q = \text{const}.$$

En el espacio H de las funciones reticulares, dadas sobre $\bar{\omega}^*$, definiremos el producto escalar

$$(u, v) = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2-1} \frac{\hbar_1(i)\,\hbar_2(j)}{\rho(i)} u(i, j)\, v(i, j). \tag{27}$$

Definamos los operadores A_1 y A_2, que actúan en H, de la siguiente manera: $A_\alpha y = -\Lambda_\alpha \bar{y}$, donde $y(i, j) = \bar{y}(i, j)$ para $0 \leqslant i \leqslant N_1$, $0 \leqslant j \leqslant N_2 - 1$ y $\bar{y}$ satisface las condiciones de periodicidad (9). Entonces el esquema (26) puede ser escrito en la forma de la ecuación operacional

$$Au = \bar{f}, \quad A = A_1 + A_2. \tag{28}$$

en el espacio H.

Para resolver la ecuación (28) utilicemos el método de las direcciones variables, cuyo esquema iterativo tiene la forma

$$\begin{aligned} &B_{k+1} \frac{y_{k+1} - y_k}{\tau_{k+1}} + A y_k = \bar{f}, \quad k = 0, 1, \ldots, y_0 \in H, \\ &B_k = (\omega_k^{(1)} E + A_1)(\omega_k^{(2)} E + A_2), \ \tau_k = \omega_k^{(1)} + \omega_k^{(2)}. \end{aligned} \tag{29}$$

La autoconjugación de los operadores A_1 y A_2 en el espacio H se establece con ayuda de la fórmula de Green de diferencias, y su conmutatividad se comprueba directamente.

Hallemos ahora las cotas de los operadores A_1 y A_2, es decir, la constantes δ_α y Δ_α, $\alpha = 1, 2$, en las desigualdades

$$\delta_\alpha (u, u) \leqslant (A_\alpha u, u) \leqslant \Delta\alpha (u, u).$$

Hallaremos primeramente δ_2 y Δ_2. Ya que para la función $\bar{u}(i, j)$, la cual satisface la condición de periodicidad (9), tenemos

$$(\Lambda_2 \bar{u}, \bar{u}) = -(\Lambda_2 \bar{u}, \bar{u}) = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2-1} \frac{\hbar_1(i)\,\hbar_2(i)}{\rho(i)} (a_2 \bar{u}^2_{\bar{\varphi}})_{ij},$$

entonces

$$\delta_2 = 0, \quad \Delta_2 = \max_{0 \leqslant j \leqslant N_2 - 1} \left[\frac{a_2(j+1)}{h_2(j+1)} + \frac{a_2(j)}{h_2(j)} \right] \frac{2}{h_2(j)}.$$

Aquí fueron tenidas en cuenta las relaciones (9) para a_2 y h_2.

A continuación, utilizando un análogo del lema 16 del capítulo V, hallaremos, que δ_1 se puede estimar de la siguiente forma: $1/b_1 = \max_{0 \leqslant i \leqslant N_1} v(i)$, donde $v(i)$ es la solución del problema de contorno

$$\begin{aligned}
&\rho(a_1 v_{\bar{r}})_{\hat{r}} - d\rho^2 v = -1, \quad 1 \leqslant i \leqslant N_1 - 1, \\
&\frac{\rho a_1^{+1}}{h_1} v_r - \left(d + \frac{r\varkappa_1^-}{\rho h_1} \right) \rho^2 v = -1, \quad i = 0, \\
&-\frac{\rho a_1}{h_1} v_{\bar{r}} - \left(d + \frac{r\varkappa_1^+}{\rho h_1} \right) \rho^2 v = -1, \quad i = N_1.
\end{aligned} \tag{30}$$

El problema (30) se resuelve por el método de factorización.

Obtengamos ahora una estimación para Δ_1. Utilizando la primera fórmula de Green de diferencias y la definición (27) para el producto escalar, hallaremos

$$(A_1 u,\ u) = -(\overline{\Lambda}_1 \overline{u},\ \overline{v}) = \sum_{j=0}^{N_2 - 1} h_2(j) \Big[\sum_{i=1}^{N_1} h_1(i) a_1(i) \overline{u}_{\bar{r}}^2(i,\ j) + d \sum_{i=0}^{N_1} h_1(i) \rho(i) \overline{u}^2(i,\ j) + l_1 \varkappa_1^- \overline{u}^2(0,\ j) + L_1 \varkappa_1^+ \overline{u}^2(N_1,\ j) \Big].$$

Estimemos la expresión que se encuentra en los corchetes. Del análogo del lema 16 del capítulo V obtendremos la estimación

$$d \sum_{i=0}^{N_1} h_1(i) \rho(i) \overline{u}^2(i,\ j) + l_1 \varkappa_1^- \overline{u}^2(0,\ j) + L_1 \varkappa_1^+ \overline{u}^2(N_1,\ j) \leqslant m_1 \Big[\sum_{j=1}^{N_1} (a_1 \overline{u}_{\bar{r}}^2)_{ij} h_1(i) + \sum_{i=0}^{N_1} \frac{h_1(i)}{\rho(i)} \overline{u}^2(i,\ j) \Big], \tag{31}$$

donde $m_1 = \max_{0 \leqslant i \leqslant N_1} w(i)$, y $w(i)$ es la solución del problema

$$\begin{aligned}
&\rho(a_1 w_{\bar{r}})_{\hat{r}} - w = -d\rho^2, \quad 1 \leqslant i \leqslant N_1 - 1, \\
&\frac{\rho a_1^{+1}}{h_1} w_r - w = -\left(d + \frac{r\varkappa_1^-}{\rho h_1} \right) \rho^2, \quad i = 0, \\
&-\frac{\rho a_1}{h_1} w_{\bar{r}} - w = -\left(d + \frac{r\varkappa_1^+}{\rho h_1} \right) \rho^2, \quad i = N_1.
\end{aligned} \tag{32}$$

A continuación, del análogo del lema 17 del capítulo V obtenemos la estimación

$$\sum_{i=1}^{N_1} a_1(i) \overline{u}_{\bar{r}}^2(i,\ j) h_1(i) \leqslant m_2 \sum_{i=0}^{N_1} \frac{h_1(i)}{\rho(i)} \overline{u}^2(i,\ j), \tag{33}$$

donde

$$m_3 = \max\left(\frac{a_1(N_1)\rho(N_1)}{\hbar_1^2(N_1)}, \frac{a_1(1)\rho(0)}{\hbar_1^2(0)}, \max_{1\leqslant i\leqslant N_1-1}\frac{2\rho(i)}{\hbar_1(i)}\left[\frac{a_1(i)}{\hbar_1(i)}+\frac{a_1(i+1)}{\hbar_1(i+1)}\right]\right).$$

De (31) y (33) se deduce la estimación

$$\sum_{i=1}^{N_1} h_1 a_1 \bar{u}_{\bar{r}}^2 + d\sum_{i=0}^{N} \hbar_1\rho\bar{u}^2 + l_1\varkappa_1^-\bar{u}^2|_{i=0} + L_1\varkappa_1^+\bar{u}^2|_{i=N_1} \leqslant$$

$$\leqslant \Delta_1 \sum_{i=0}^{N_1}\frac{\hbar_1}{\rho}\bar{u}^2, \quad \Delta_1 = m_1 + m_3(1+m_1).$$

Multiplicando esta igualdad por $\hbar_2(j)$ y sumando según j desde 0 hasta $N_2 - 1$, obtendremos $(A_1 u, u) \leqslant \Delta_1 (u, u)$.

Así pues, hemos hallado las constantes δ_α y Δ_α, $\alpha = 1, 2$. Recordemos que las fórmulas para los parámetros iterativos $\omega_k^{(1)}$ y $\omega_k^{(2)}$ fueron obtenidas en el punto 4 del § 1 del cap. XI.

De modo análogo se construye el método de las direcciones variables para el problema de diferencias (8), (10) con las condiciones de contorno de primer género. Las constantes δ_α y Δ_α se estiman igual que en el caso examinado anteriormente, solamente en (30) y (32) las condiciones de contorno de tercer género se deben cambiar por las condiciones $v(0) = 0$, $v(N_1) = 0$ y $w(0) = 0$, $w(N_1) = 0$.

Como conclusión señalemos, que para $d = 0$, $\varkappa_i^\pm = 0$ el problema (8), (11), (12) es degenerado, y si se cumple la condición de solubilidad

$$\sum_{i=0}^{N_1}\sum_{j=0}^{N_2-1}\psi\rho\hbar_1\hbar_2 + \sum_{j=0}^{N_2-1}\hbar_2[L_1 g_1^+ + l_1 g_1^-] = 0,$$

entonces el problema tiene solución no única. Para este caso la colección de los parámetros $\omega_k^{(1)}$ y $\omega_k^{(2)}$ para el método de las direcciones variables (29) está construída en el punto 1 del § 4 del cap. XII.

6. **Resolución de problemas de diferencias en un sector anular.** Examinemos los métodos de resolución de los problemas de contorno de diferencias para una ecuación elíptica sin derivadas mixtas, prefijada en un sector anular.

En la región $\bar{G} = \{l_1 \leqslant r \leqslant L_1,\ l_2 \leqslant \varphi \leqslant L_2,\ l_1 > 0,\ L_2 - l_2 < 2\pi\}$ se exige hallar la solución de las ecua-

ciones (1), que satisfaga sobre los lados $r = l_1$ y $r = L_1$ una de las condiciones de contorno (2)—(5), y sobre los lados $\varphi = l_2$ y $\varphi = L_2$ una de las condiciones

$$u(r, \varphi) = g_2^-(r), \quad \varphi = l_2 \tag{34}$$

ó

$$\frac{k_2}{r}\frac{\partial u}{\partial \varphi} = \varkappa_2^- u - g_2^+(r), \quad \varphi = l_2, \tag{35}$$

$$u(r, \varphi) = g_2^+(r), \quad \varphi = L_2, \tag{36}$$

ó

$$-\frac{k_2}{r}\frac{\partial u}{\partial \varphi} = \varkappa_2^+ u - g_2^+(r), \quad \varphi = L_2. \tag{37}$$

Se supone que los coeficientes satisfacen las condiciones

$k_1(r, \varphi) \geqslant c_1 > 0,$

$k_2(r, \varphi) \geqslant c_1 > 0, \quad q(r, \varphi) \geqslant 0, \quad \varkappa_1^{\pm}(\varphi) \geqslant 0,$

$\varkappa_2^{\pm}(r) \geqslant 0.$

En la región $\bar{G}$ se introduce una red rectangular $\bar{\omega}$ no uniforme arbitraria (véase el punto 1 del § 3):

$\bar{\omega} = \{(r_i, \varphi_j) \in \bar{G},\ r_i = r_{i-1} + h_1(i),\ 1 \leqslant i \leqslant N_1,$

$r_0 = l_1, \quad r_{N_1} = L_1, \quad \varphi_j = \varphi_{j-1} + h_2(j), \quad 1 \leqslant j \leqslant N_2,$

$\varphi_0 = l_2, \quad \varphi_{N_2} = L_2\}$

y se definen los pasos medios $\hbar_1(i)$ y $\hbar_2(j)$:

$$\hbar_\alpha(m) = \begin{cases} 0{,}5h_\alpha(1), & m = 0, \\ 0{,}5\,[h_\alpha(m) + h_\alpha(m+1)], & 1 \leqslant m \leqslant N_\alpha - 1, \\ 0{,}5h_\alpha(N_\alpha), & m = N_\alpha, \end{cases} \quad \alpha = 1, 2.$$

La ecuación (1) se aproxima por la ecuación en diferencias

$$\frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} + \frac{1}{\rho^2}(a_2 y_{\bar{\varphi}})_{\hat{\varphi}} - dy = -\psi, \tag{38}$$

$$1 \leqslant i \leqslant N_1 - 1, \quad 1 \leqslant j \leqslant N_2 - 1.$$

Las condiciones de contorno de primer género (2), (4), (34), (36) se aproximan exactamente:

$$y(N_1, j) = g_1^+(\varphi_j), \quad y(0, j) = g_1^-(\varphi_j), \quad 0 \leqslant j \leqslant N_2, \tag{39}$$

$$y(i, N_2) = g_2^+(r_i), \quad y(i, 0) = g_2^-(r_i), \quad 0 \leqslant i \leqslant N_1. \tag{40}$$

Las condiciones de tercer género (3) y (5), prefijadas para $r = L_1$ y $r = l_1$, se cambian para $1 \leqslant j \leqslant N_2 - 1$ por las condiciones (11) y (12).

El análogo de diferencias de las condiciones de contorno (35) y (37) tiene la forma

$$\frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} + \frac{a_2^{+1}}{\rho^2 h_2} y_{\bar{\varphi}} - \left(d + \frac{\varkappa_2^-}{\rho h_2}\right) y = -\psi - \frac{g_2^-}{\rho h_2}, \quad j = 0, \qquad (41)$$

$$\frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} - \frac{a_2}{\rho^2 h_2} y_{\bar{\varphi}} - \left(d + \frac{\varkappa_2^+}{\rho h_2}\right) y = -\psi - \frac{g_2^+}{\rho h_2}, \quad j = N_2. \qquad (42)$$

Si sobre los lados del rectángulo que se intersecan están dadas condiciones de contorno de tercer género, entonces en los nodos angulares de la red $\bar{\omega}$ se ponen las siguientes condiciones de contorno:

$$\frac{a_1^{+1}}{\rho_1 h_1} y_r + \frac{a_2^{+1}}{\rho^2 h_2} y_{\bar{\varphi}} - \left(d + \frac{r\varkappa_1^-}{\rho h_1} + \frac{\varkappa_2^-}{\rho h_2}\right) y = -\psi - \frac{r g_1^-}{\rho h_1} - \frac{g_2^-}{\rho h_2}, \qquad (43)$$

si $i = j = 0$;

$$-\frac{a_1}{\rho h_1} y_{\bar{r}} + \frac{a_2^{+1}}{\rho^2 h_2} y_{\bar{\varphi}} - \left(d + \frac{r\varkappa_1^+}{\rho h_1} + \frac{\varkappa_2^-}{\rho h_2}\right) y = -\psi - \frac{r g_1^+}{\rho h_1} - \frac{g_2^-}{\rho h_2}, \qquad (44)$$

si $i = N_1$, $j = 0$;

$$\frac{a_1^{+1}}{\rho h_1} y_r - \frac{a_2}{\rho^2 h_1} y_{\bar{\varphi}} - \left(d + \frac{r\varkappa_1^-}{\rho h_1} + \frac{\varkappa_2^+}{\rho h_2}\right) y = -\psi - \frac{r g_1^-}{\rho h_1} - \frac{g_2^+}{\rho h_2}, \qquad (45)$$

si $i = 0$, $j = N_2$; y, finalmente,

$$-\frac{a_1}{\rho h_1} y_{\bar{r}} - \frac{a_2}{\rho^2 h_2} y_{\bar{\varphi}} - \left(d + \frac{r\varkappa_1^+}{\rho h_1} + \frac{\varkappa_2^+}{\rho h_2}\right) y = -\psi - \frac{r g_1^+}{\rho h_1} - \frac{g_2^+}{\rho h_2}, \qquad (46)$$

si $i = N_1$, $j = N_2$.

Si se examina el problema en diferencias (38), (11), (12), (41)—(46) con $d \equiv 0$, $\varkappa_i^{\pm} \equiv 0$, $\alpha = 1, 2$, entonces la solución existe, si se cumple la condición

$$\sum_{j=0}^{N_2}\sum_{i=0}^{N_1} \rho h_1 h_2 \psi + \sum_{j=0}^{N_2} h_2 (L_1 g_1^+ + l_1 g_1^-) + \sum_{i=0}^{N_1} h_1 (g_2^- + g_2^+) = 0,$$

la cual es el análogo de diferencias de la condición (27) del § 1 de solubilidad del problema correspondiente para la ecuación diferencial. Con esto dos soluciones cualesquiera del problema indicado se diferencian en una constante.

La afirmación citada se demuestra casi igual que como fue hecho en el punto 2 del § 3, para el caso de un círculo y un anillo. Aquí el producto escalar en el espacio H de las funciones reticulares, dadas sobre $\overline{\omega}$, se define por la fórmula

$$(u, v) = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} u(i, j)\, v(i, j)\, \rho(i)\, \hbar_1(i)\, \hbar_2(j). \qquad (47)$$

Señalemos, que los coeficientes a_1, a_2, q y la función $\rho(i)$ en este punto se determinan, como en el punto 1 del § 3.

Hagamos una observación con respecto a los métodos de resolución de los problemas de diferencias construidos. Si los coeficientes k_1, k_2 y q dependen de r, $\varkappa_i^{\pm}$ son las constantes, y $\varkappa_s^{\pm} = 0$, si están dadas las condiciones de contorno (3), (5), (35), (37), y si la red $\overline{\omega}$ es uniforme según φ, entonces los problemas de contorno correspondientes pueden ser resueltos por los métodos directos, construídos en los capítulos III y IV.

Si se cumplen las condiciones $k_1 = k_1(r)$, $k_2 = k_2(\varphi)$, $q = \text{const}$, $\varkappa_\alpha^{\pm} = \text{const}$ y la red $\overline{\omega}$ es no uniforme por cada una de las direcciones, entonces para la resolución de los problemas de diferencias se puede utilizar el método de las direcciones variables con el conjunto óptimo de los parámetros. En este caso, al igual que fue hecho en el punto anterior, las ecuaciones en diferencias se deben multiplicar previamente por $\rho^2(i)$.

7. Caso general de coeficientes variables. Examinemos ahora el caso, cuando las variables no se separan y la solución del problema de contorno de diferencias se encuentra con un método iterativo.

Supongamos, por ejemplo, que se exige hallar la solución del problema de Dirichlet para la ecuación (1) sobre la red $\overline{\omega}$ bajo las suposiciones, de que la red ω es uniforme según $\varphi(h_2(j) \equiv h_2)$, $q = 0$, y los coeficientes k_1 y k_2 satisfacen las condiciones

$$0 < c_1 \leqslant k_\alpha(r, \varphi) \leqslant c_2, \quad \alpha = 1, 2, \qquad (48)$$

Bajo estas suposiciones el problema de diferencias se escribe en la forma

$$\Lambda y = \frac{1}{\rho}(a_1 y_{\bar{r}})_{\hat{r}} + \frac{1}{\rho^2}(a_2 y_{\bar{\varphi}})_{\varphi} = -\psi, \ (r, \varphi) \in \omega,$$
$$y(r, \varphi) = g(r, \varphi), \quad (r, \varphi) \in \gamma, \tag{49}$$

donde

$$a_1(i, j) = \bar{r}_i k_1(\bar{r}_i, \varphi_j), \quad a_2(i, j) = k_2(r_i, \varphi_j),$$
$$\bar{r}_i = r_i - 0{,}5h_1(i), \quad \bar{\varphi}_j = \varphi_j - 0{,}5h_2. \tag{50}$$

En el espacio H de las funciones reticulares, dadas sobre ω, definiremos el producto escalar

$$(u, v) = \sum_{i=1}^{N_1-1} \sum_{j=1}^{N_2-1} u(i, j) v(i, j) \rho(i) \hbar_1(i) h_2$$

y los operadores A y R, los cuales actúan en H, $Ay = \Lambda \overset{\circ}{y}$, $Ry = -\mathcal{R}\overset{\circ}{y}$, donde $y(r, \varphi) = \overset{\circ}{y}(r, \varphi)$ para $(r, \varphi) \in \omega$ y $\overset{\circ}{y}(r, \varphi) = 0$ para $(r, \varphi) \in \gamma$. Aquí el operador en diferencias $\mathcal{R}$ se define por la relación

$$\mathcal{R}y = \frac{1}{\rho}(\bar{r} y_{\bar{r}})_{\hat{r}} + \frac{1}{\rho^2} y_{\bar{\varphi}\varphi}, \quad (r, \varphi) \in \omega.$$

Utilizando las fórmulas de Green de diferencias, se puede verificar, que los operadores A y R son autoconjugados en H y, además, para cualquier $y \in H$ tienen lugar las igualdades

$$(Ay, y) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2-1} a_1 \overset{\circ}{y}_{\bar{r}}^2 h_1 h_2 + \sum_{j=1}^{N_2} \sum_{i=1}^{N_1-1} \frac{a_2}{\rho} \overset{\circ}{y}_{\bar{\varphi}}^2 \hbar_1 h_2,$$
$$(Ry, y) = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2-1} \bar{r} \overset{\circ}{y}_{\bar{r}}^2 h_1 h_2 + \sum_{j=1}^{N_2} \sum_{i=1}^{N_1-1} \frac{1}{\rho} \overset{\circ}{y}_{\bar{\varphi}}^2 \hbar_1 h_2.$$

De aquí y de (48), (50) se deduce, que los operadores A y R son energéticamente equivalentes con constantes $\gamma_1 = c_1$ y $\gamma_2 = c_2$:

$$\gamma_1 (Ry, y) \leqslant (Ay, y) \leqslant \gamma_2 (Ry, y), \ \gamma_1 > 0. \tag{51}$$

El problema de diferencias (49) puede ser escrito en la forma de la ecuación operacional

$$Au = f$$

con el operador A definido más arriba. Para su resolución utilicemos el esquema iterativo implícito

$$B\frac{y_{k+1}-y_k}{\tau_{k+1}}+Ay_k=f,\quad k=0,\ 1,\ \ldots,\ y_0\in H, \tag{52}$$

donde $B = R$.

De la teoría general de los métodos iterativos, expuestos en el capítulo VI, se deduce que si los parámetros τ_{k+1} en el esquema (52) se eligen mediante las fórmulas del método de Chebishev

$$\tau_k=\frac{\tau_0}{1+\rho_0\mu_k},$$

$$\mu_k\in\mathfrak{M}_n^*=\left\{-\cos\frac{(2i-1)\,\pi}{2n},\ 1\leqslant i\leqslant n\right\},\quad k=1,\ 2,\ \ldots,\ n,$$

entonces para el error $z_n = y_n - u$ será cierta la estimación

$$\|y_n-u\|_D\leqslant\varepsilon\,\|y_0-u\|_D,$$

donde $D = A$ ó $D = B$, $D = AB^{-1}A$, y el número de iteraciones satisface la estimación

$$n\geqslant n_0(\varepsilon)=\ln(0{,}5\varepsilon)/\ln\rho_1.$$

Aquí

$$\tau_0=\frac{2}{\gamma_1+\gamma_2},\quad \rho_0=\frac{1-\xi}{1+\xi},\quad \rho_1=\frac{1-\sqrt{\xi}}{1+\sqrt{\xi}},\quad \xi=\frac{\gamma_1}{\gamma_2}.$$

Ya que γ_1 y γ_2 no dependen de los pasos de la red $\bar{\omega}$, entonces el número de iteraciones es proporcional a $|\ln 0{,}5\varepsilon|$ y no cambia al hacerse más fina la red.

Para encontrar y_{k+1} obtendremos el problema de diferencias

$$\mathscr{R}y_{k+1}=-F,\quad (r,\ \varphi)\in\omega,\quad y_{k+1}=g,\quad (r,\ \varphi)\in\gamma$$

del segundo miembro conocido $F=-\mathscr{R}y_k+\tau_{k+1}(\Lambda y_k+\psi)$. Señalemos que este problema satisface todas las condiciones, que permiten encontrar su solución por uno de los métodos directos, por ejemplo, por el método de reducción completa con un gasto de $O(N_1N_2\log_2 N_2)$ operaciones aritméticas para $N_2 = 2^n$. De esta forma, el número total de operaciones que es necesario gastar para encontrar la solución del problema de diferencias examinado con exactitud ε, se estima por una magnitud $O(N_1N_2\log_2 N_2\ln(2/\varepsilon))$.

De manera análoga, pueden ser construídos, bajo las suposiciones correspondientes, los métodos iterativos de resolución de los problemas de contorno de diferencias planteados en los párrafos anteriores en sistemas cilíndricos y polares de coordenadas.

Construcción del polinomio de menor desviación del cero

1. En el § 2 del cap. VI durante el examen de los esquemas iterativos de dos capas fue formulado un problema: construir un polinomio de grado n, que tome el valor 1 en cero, y cuyo módulo máximo sobre el segmento $[\gamma_1, \gamma_2]$ sea mínimo.

Resolvamos este problema. Nos será cómodo realizar todas las investigaciones no sobre el segmento $[\gamma_1, \gamma_2]$, sino sobre el segmento $[-1, 1]$. Para esto haremos una sustitución lineal de variable, que lleve el segmento $\gamma_1 \leqslant t \leqslant \gamma_2$ al segmento $-1 \leqslant x \leqslant 1$, y el punto γ_1 al punto 1. Este cambio tiene la forma

$$t = \frac{1-\rho_0 x}{\tau_0}, \quad \tau_0 = \frac{2}{\gamma_1+\gamma_2}, \quad \rho_0 = \frac{1-\xi}{1+\xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Para tal cambio al punto $t = 0$ le corresponde el punto $x = 1/\rho_0 > 1$.

De esta manera, el problema formulado más arriba es equivalente al problema: entre todos los polinomios de grado n que toman el valor 1 en el punto $x = 1/\rho_0 > 1$, hallar el de menor desviación del cero en el segmento $[-1, 1]$.

Este es el problema clásico de Chebishev de la teoría de aproximación de funciones, cuya solución es conocida, pero a nosotros nos será útil hallar esta solución de nuevo. Para esto tendremos necesidad de:

TEOREMA 1. *Cualesquiera sean las funciones continuas $g(x) > 0$ y $f(x)$ sobre $[-1, 1]$, existe un único polinomio $P_n(x)$ de grado no superior de n tal que*

$$q_n = \max_{-1 \leqslant x \leqslant 1} g(x) |f(x) - P_n(x)| = \min_{\{R_k(x)\}_{k \leqslant n}} \max_{-1 \leqslant x \leqslant 1} g(x) |f(x) - R_k(x)|.$$

Este polinomio se caracteriza totalmente por la propiedad siguiente: el número de puntos sucesivos del segmento $[-1, 1]$, en los cuales la función $g(x)(f(x) - P_n(x))$ toma el valor q_n con signos alternados no es menor que $n + 2$.

Transformemos el problema planteado al problema que figura en el teorema 1. Teniendo en cuenta, que el polinomio buscado toma el

valor 1 en el punto $x = 1/\rho_0$, representémoslo en la forma

$$P_n(x) = 1 - \left(\frac{1}{\rho_0} - x\right) R_{n-1}(x) = \frac{1-\rho_0 x}{\rho_0}\left[\frac{\rho_0}{1-\rho_0 x} - R_{n-1}(x)\right],$$

donde $R_{n-1}(x)$ es un polinomio de grado no superior de $n-1$.

De aquí se deduce que nuestro problema se reduce al problema de buscar el polinomio $R_{n-1}(x)$ de grado no mayor que $n-1$, el cual da la mejor aproximación uniforme con peso $g(x) = (1-\rho_0 x)/\rho_0 > 0$ de la función $f(x) = \rho_0/(1-\rho_0 x)$ sobre el segmento $[-1, 1]$.

Precisamente este problema figura en el teorema 1.

Por eso sobre la base del teorema 1 existen al menos $n+1$ puntos $x_1, x_2, \ldots, x_{n+1}$ del segmento $[-1, 1]$, en los cuales el polinomio $P_n(x)$ buscado toma el valor q_n con signos alternados.

Al principio mostremos que tales puntos deben ser exactamente igual a $n+1$. En efecto, para que una función continua pueda tomar los valores diferentes de cero q_n con signos alternados en más $n+1$ puntos sucesivos del segmento $[-1, 1]$, ella debe anularse sobre este segmento en no menos de n puntos.

Puesto que el polinomio $P_n(x)$ es diferente del idénticamente nulo, entonces sobre el segmento $[-1, 1]$ él puede anularse en no más de n puntos. Por eso mismo, el polinomio $P_n(x)$ buscado toma el valor q_n con signos alternados exactamente $n+1$ veces sobre $[-1, 1]$.

Caractericemos estos puntos. Si el polinomio $P_n(x)$ toma el valor máximo en un punto interior del segmento $[-1, 1]$ entonces la derivada $P_n'(x)$ se anula en este punto. Pero el grado de $P_n'(x)$ es igual a $n-1$ y, por consiguiente, la derivada del polinomio buscado puede anularse solamente en $n-1$ puntos. Por eso el polinomio buscado tiene $n-1$ puntos extremales interiores sobre $[-1, 1]$ y por lo tanto, dos extremos de frontera, es decir

$$|P_n(-1)| = |P_n(1)| = q_n.$$

Así pues, tenemos

$$P_n(\omega_j) = 0, \quad j = 1, 2, \ldots, n, \quad |P_n(x_j)| = q_n, \quad j = 1, 2, \ldots, n+1,$$

donde ω_j son las raíces del polinomio y x_j los puntos extremales

$$-1 = x_{n+1} < \omega_n < x_n < \ldots < \omega_2 < x_2 < \omega_1 < x_1 = 1,$$

Además, como $P_n(1/\rho_0) = 1$ y todas las raíces del polinomio $P_n(x)$ descansan sobre el segmento $[-1, 1]$, entonces $P_n(1) = q_n$ y, por consiguiente, son válidas las igualdades

$$P_n(x_j) = (-1)^{j-1} q_n, \quad j = 1, 2, \ldots, n+1. \tag{1}$$

Tiene lugar

LEMA 1. *El polinomio $P_n(x)$, el cual entre todos los polinomios de n-ésimo grado que toman el valor 1 para $x = 1/\rho_0$, es el de menor desviación del cero sobre el segmento $[-1, 1]$, satisface la ecuación diferencial*

$$(1-x^2)(P')^2 = n^2(q_n^2 - P^2). \tag{2}$$

Realmente, por lo demostrado los puntos $x_2, x_3, \ldots, x_n$ son ceros simples del polinomio $P_n'(x)$. Es obvio, que estos puntos son ceros de multiplicidad dos del polinomio $q_n^2 - P_n^2(x)$, y por lo demostrado

los puntos $x_{n+1} = -1$ y $x_1 = 1$ son ceros simples de este polinomio. Por eso los polinomios $(1 - x^2)(P_n'(x))^2$ y $q_n^2 - P_n^2(x)$ de grado 2 tienen los mismos ceros. Por lo tanto, ellos son proporcionales, es decir

$$(1-x^2)(P_n')^2 = c(q_n^2 - P_n^2(x)).$$

Igualando los coeficientes de mayor grado según x de ambos polinomios, encontramos $c = n^2$. El lema está demostrado.

2. Pasemos a la construcción del polinomio $P_n(x)$, utilizando la ecuación (2). Esta ecuación, además de la función desconocida $P_n(x)$, contiene también el parámetro desconocido q_n. No fijaremos por separado condiciones complementarias, que definan univocamente la solución de la ecuación (2), sino utilizaremos toda la información conocida con respecto a $P_n(x)$.

Al principio examinemos la ecuación (2) sobre el segmento $[-1, 1]$. En este caso $|P_n(x)| \leqslant q_n$, y, por consiguiente, de los primer y segundo miembros de la ecuación (2) se puede extraer la raíz

$$\pm \frac{dP}{\sqrt{q_n^2 - P^2}} = n \frac{dx}{\sqrt{1-x^2}}, \quad 0 \leqslant x \leqslant 1. \tag{3}$$

Investiguemos el primer miembro de (3). Si $P_n(x_{j+1}) = q_n$, entonces al cambiar x desde x_{j+1} hasta x_j la función $P_n(x)$ decrece desde q_n hasta $-q_n$. Con esto el diferencial dP es negativo, y por eso en el primer miembro de la ecuación (3) se debe tomar el signo menos. Análogamente encontramos, que si $P_n(x_{j+1}) = -q_n$, entonces se debe tomar el signo más. Tomando en consideración (1), obtendremos, que sobre el segmento $[x_{j+1}, x_j]$ la ecuación (3) debe ser escrita de manera

$$(-1)^{j-1} \frac{dP}{\sqrt{q_n^2 - P^2}} = n \frac{dx}{\sqrt{1-x^2}}, \quad x \in [x_{j+1}, x_j],$$

$$j = 1, 2, \ldots, n. \tag{4}$$

Obtengamos ahora la expresión para $P_n(x)$ sobre el segmento $[-1, 1]$. Sea x cualquier punto del segmento $[-1, 1]$, y para definición supongamos que x pertenece, por ejemplo, al segmento $[x_{k+1}, x_k]$.

Integremos el segundo miembro de la ecuación (4) según x desde x hasta 1. Obtendremos

$$n \int_x^1 \frac{dx}{\sqrt{1-x^2}} = n \arcsen x \Big|_x^1 = n \arccos x.$$

Integremos el primer miembro de la ecuación (4). Cuando x cambia desde x_{j+1} hasta x_j, la función $P(x)$ cambia de $P(x_{j+1}) = (-1)^j q_n$ hasta $P(x_j) = (-1)^{j-1} q_n$. Por eso

$$(-1)^{j-1} \int_{P(x_{j+1})}^{P(x_j)} \frac{dP}{\sqrt{q_n^2 - P^2}} = \int_{-q_n}^{q_n} \frac{dP}{\sqrt{q_n^2 - P^2}} =$$

$$= \arcsen \frac{P}{q_n} \Big|_{-q_n}^{q_n} = \pi.$$

A continuación, al integrar el primer miembro de (4) desde $P(x)$ hasta $P(x_k)$ obtendremos

$$(-1)^{k-1}\int\limits_{P(x)}^{P(x_k)}\frac{dP}{\sqrt{q_n^2-P^2}}=\int\limits_{(-1)^{k-1}P(x)}^{q_n}\frac{dP}{\sqrt{q_n^2-P}}=$$

$$=\arccos(-1)^{k-1}\frac{P(x)}{q_n}.$$

Puesto que

$$\int\limits_x^1\frac{dx}{\sqrt{1-x^2}}=\int\limits_x^{x_k}\frac{dx}{\sqrt{1-x^2}}+\sum_{j=1}^{k-1}\int\limits_{x_{j+1}}^{x_j}\frac{dx}{\sqrt{1-x^2}},$$

entonces definitivamente obtenemos

$$n\arccos x=(k-1)\pi+\arccos(-1)^{k-1}\frac{P(x)}{q_n}. \tag{5}$$

De aquí hallaremos

$$P_n(x)=q_n\cos(n\arccos x),\quad |x|\leqslant 1. \tag{6}$$

Suponiendo en (5) $x=\omega_k\in[x_{k+1},\ x_k]$, hallaremos las raíces del polinomio $P_n(x)$

$$\omega_k=\cos\frac{(2k-1)\pi}{2n},\quad k=1,\ 2,\ \ldots,\ n.$$

La fórmula (6) define el polinomio $P_n(x)$ para $x\in[-1,\ 1]$. Hallemos la forma del polinomio $P_n(x)$ para $|x|\geqslant 1$ y determinemos q_n. Para esto notemos que

$$\omega_{n-k+1}=\cos\left(\pi-\frac{2k-1}{2n}\pi\right)=-\omega_k,\quad k=1,\ 2,\ \ldots,\ n.$$

Por eso $P_n(-x)=(-1)^nP_n(x)$ y, por lo tanto, es suficiente determinar $P_n(x)$ para $x\geqslant 1$.

Investiguemos la ecuación (2) para $x\geqslant 1$. En este caso ella se debe volver a escribir de la forma siguiente:

$$(x^2-1)(P')^2=n^2(P^2-q_n^2),\quad x\geqslant 1.$$

Puesto que $x\geqslant 1$, entonces $P(x)\geqslant q_n$ y la función crece. Por eso, extrayendo la raíz, obtenemos

$$\frac{dP}{\sqrt{P^2-q_n^2}}=n\frac{dx}{\sqrt{x^2-1}}.$$

Al integrar el miembro derecho de esta ecuación desde 1 hasta x el miembro derecho se integrará desde q_n hasta $P_n(x)$. Por eso

$$\int\limits_{q_n}^{P_n(x)}\frac{dP}{\sqrt{P^2-q_n^2}}=\ln\left(\frac{P_n(x)}{q_n}+\sqrt{\frac{P_n^2(x)}{q_n^2}-1}\right)=$$

$$=\operatorname{arcch}\frac{P_n(x)}{q_n}=n\int\limits_1^x\frac{dx}{\sqrt{x^2-1}}=n\ln(x+\sqrt{x^2-1})=n\operatorname{arcch}x. \tag{7}$$

De aquí obtenemos

$$P_n(x) = q_n \operatorname{ch}(n \operatorname{arcch} x), \quad x \geqslant 1.$$

Ya que $P_n(x) = (-1)^n P_n(-x)$, entonces para $x \leqslant 1$ hallaremos

$$P_n(x) = (-1)^n q_n \operatorname{ch}(n \operatorname{arcch}(-x)) = q_n \operatorname{ch}(n \operatorname{arcch} x), \quad x \leqslant -1.$$

De este modo, para $|x| \geqslant 1$ obtendremos la siguiente expresión para el polinomio $P_n(x)$:

$$P_n(x) = q_n \operatorname{ch}(n \operatorname{arcch} x), \quad |x| \geqslant 1. \tag{8}$$

Hallemos ahora q_n. Poniendo en (8) $x = 1/\rho_0$, y teniendo en cuenta, que $P_n(1/\rho_0) = 1$, obtendremos

$$q_n = 1/\operatorname{ch}(n \operatorname{arcch}(1/\rho_0)).$$

Por otra parte, poniendo en (7) $x = 1/\rho_0$, hallaremos

$$\ln \frac{1+\sqrt{1-q_n^2}}{q_n} = n \ln \frac{1+\sqrt{1-\rho_0^2}}{\rho_0} = n \ln \frac{1}{\rho_1},$$

donde

$$\rho_1 = \frac{\rho_0}{1+\sqrt{1-\rho_0^2}} = \frac{1-\sqrt{\xi}}{1+\sqrt{\xi}}, \quad \xi = \frac{\gamma_1}{\gamma_2}, \quad \rho_0 = \frac{2\rho_1}{1+\rho_1^2}.$$

Por lo tanto

$$q_n = \frac{1}{\operatorname{ch}\left(n \operatorname{arcch} \dfrac{1}{\rho_0}\right)} = \frac{2\rho_1^n}{1+\rho_1^{2n}} < 1 \tag{9}$$

Uniendo (6) y (8) obtendremos

$$P_n(x) = q_n T_n(x) = T_n(x)/T_n(1/\rho_0), \tag{10}$$

donde

$$T_n(x) = \begin{cases} \cos(n \arccos x), & |x| \leqslant 1, \\ \operatorname{ch}(n \operatorname{arcch} x), & |x| \geqslant 1. \end{cases}$$

El polinomio $T_n(x)$ se llama polinomio de Chebishev de primer género y de grado n.

Así pues, el problema planteado está totalmente resuelto. Su solución se da mediante las fórmulas (9), (10). Regresando a la variable t, obtendremos el polinomio buscado

$$Q_n(t) = P_n\left(\frac{1-\tau_0 t}{\rho_0}\right) = q_n T_n\left(\frac{1-\tau_0 t}{\rho_0}\right),$$

que es el de menor desviación del cero sobre el segmento $[\gamma_1, \gamma_2]$.

Bibliografía

1. *Wasow W., Forsythe G.*, Finite-Difference Methods for Partial Differential Equations, N.-Y., 1960.
2. *Гельфонд А. О.* Исчисление конечных разностей.— М.: Наука, 1967. (Guelfond A. O., Cálculo de diferencias finitas).
3. *Марчевский М. М., Ляшко А. Д.* Разностные схемы для нелинейных задач математической физики.— Казань: Ротапринт, изд. Каз. гос. ун., 1976.
(*Karchevski M. M. Liachko A. D.*, Esquemas de diferencias para problemas no lineales de la física matemática)
4. *Красносельский М. А., Вайникко Г. М.* и др. Приближенное решение операторных уравнений.— М.: Наука, 1969.
(*Krasnoselski M. A., Vainikko G. M.* y otros, Resolución aproximada de ecuaciones operacionales)
5. *Марчук Г. И.* Методы вычислительной математики.— Новосибирск: Наука, 1973.
(*Marchuk G. I.*, Métodos de la matemática de cálculo)
6. *Оганесян Л. А., Ривкинд В. Я., Руховец Л. А.* Вариационно-разностные методы решения эллиптических уравнений, ч. 1 и 2.— В сб.: Дифференциальные уравнения и их применение, вып. 5, Вильнюс, Пяргале, 1973, вып. 8, Вильнюс, Пяргале, 1974.
(*Oganesian L. A., Rivkind V. Ya., Rujovetz L. A.*, Métodos variacionales de diefrencias de ecuaciones elípticas; partes 1 y 2. En la colección: Ecuaciones diefrenciales y sus aplicaciones)
7. *Ortega J. M., Rheinboldt W. C.* Iteration solucion of non-linear equations with several variables, N.-Y., 1970.
8. *Самарский А. А.* Введение в теорию разностных схем.— М.: Наука, 1971 (имеется библиография до 1971 г.)
(*Samarski A. A.*, Introducción a la teoría de los esquemas de diferencias)
9. *Самарский А. А.* Теория разностных схем.— М.: Наука, 1977.
(*Samarski A. A.*, Teoría de esquemas de diferencias)
10. *Самарский А. А. Гулин А. В.* Устойчивость разностных схем.— М.: Наука, 1973.
(*Samarski A. A., Gulin A. V.*, Estabilidad de los esquemas de diferencias)
11. *Самарский А. А., Андреев В. Б.* Разностные методы для эллиптических уравнений. М.: Наука, 1976.
(*Samarski A. A., Andreiev V. B.*, Métodos de diferencias para las ecuaciones elípticas)
12. *Самарский А. А., Карамзин Ю. Н.*, Разностные уравнения.— М.: Знание, 1978.
(*Samarski A. A., Karamzin Yu. N.*, Ecuaciones en diferencias)
13. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры.— М.: Физматгиз, 1963.
(*Faddeev D. K., Faddeeva V. N.*, Métodos de cálculo del álgebra lineal)
14. *Yong D. M.*, Iterative solution of large linear systems: N.-Y., L.: Acad. Press, 1971.

Indice alfabético

Aceleración de la convergencia II, 129
Algoritmo de la transformación de Fourier discreta I, 209

Conjunto de parámetros de Chebishev II, 16, 29
Corrección I, 333, II, 22

Derivada de Gato I, 274, II, 305
Derivadas de diferencias I, 35

Ecuación en diferencias I, 38
— reticular I, 38
Elemento propio del operador I, 285, 286
Error II, 22
Esquema de diferencias I, 31
Esquemas iterativos del operador factorizado II, 346
Estabilidad computacional II, 22
— respecto a los datos a priori II, 84

Forma canónica de un esquema iterativo de dos capas I, 326
— — — — — — — — tres capas estándar I, 327, II, 72
Fórmulas de Green de diferencias I, 295
Función de Green del operador de diferencias I, 301
— reticular I, 32
— — vectorial I, 33

Identidades de diferencias I, 294

Método alternativo triangular II, 173, 192
— de Chebishev (Richardson) II, 15, 259
— — descenso gradiental II, 319
— — establecimiento I, 325
— — factorización (exploración) I, 95
— — — cíclica I, 110
— — — matricial I, 135
— — — no monótona I, 110
— — — ortogonal I, 142
— — iteración simple II, 34
— las correcciones conjugadas II, 123
— — — — mínimas II, 107
— — — direcciones conjugadas II, 110
— — — — variables II, 218, 244, 271
— de los defectos conjugados II, 123
— — — — mínimos II, 100, 281
— — — errores conjugados II, 124, 264
— — — — mínimos II, 109, 263
— — — gradientes conjugados II, 123
— — Newton-Kantorovich II, 310
— — reducción I, 172
— — relajación superior II, 148
— — Seidel II, 138
— — separación variable I, 241
— — variación de constantes I, 54
— del descenso más rápido II, 105
— iterativo de tres capas estacionario II, 81
— semi-iterativo de Chebishev II, 76
Métodos iterativos biotápicos II, 313, 351
— — de tipo variacional II, 94
— — triangulares II, 165

Núcleo del operador I, 276

Operador autoconjugado I, 278
— conjugado I, 277
— continuo I, 274
— de diferencias I, 34
— definido positivo I, 279
— de resolución I, 333, II, 13
— — transición I, 333, II, 13
— fuertemente monótono I, 280, II, 303
— monótono I, 279, II, 303
— normal I, 278
— potencial II, 317
Operadores conmutativos I, 275
— energéticamente equivalentes I, 280

Polinomio de Chebishev de primer género I, 73
— — — — segundo género I, 74
Principio de las aplicaciones contraídas I, 288
— — regularización II, 341
Problema en valores propios I, 80

Propiedad asintónica II, 101
Radio espectral I, 276, II, 149
— numérico del operador I, 279, II, 45
Red (retículo) I, 30
Regularizador II, 343

Solución generalizada I, 288, II, 276
— normal I, 287, II, 276

Valor propio del operador I, 285, 286
Variante por flujos del método de factorización I, 106